INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

On the Grobman-Hartman theorem for control systems

Laurent Baratchart — Monique Chyba — Jean-Baptiste Pomet

N° 5040

décembre 2003

THÈME 4

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5040--FR+ENG

On the Grobman-Hartman theorem for control systems

Laurent Baratchart* , Monique Chyba† , Jean-Baptiste Pomet‡

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet MIAOU

Rapport de recherche n° 5040 — décembre 2003 — 92 pages

Abstract: We consider the problem of topological linearization of control systems, i.e. local equivalence to a linear controllable system *via* transformations that are topological but nto necessarily differentiable. On the one hand we prove that, when point-wise transformations are considered (static feedback transformations), topological linearization implies smooth linearization, at least away from singularities. On the other hand, if we allow the transformation to depend on the control at a functional level so as to define a flow (open loop transformations), we prove a version of the Grobman-Hartman theorem for control systems.

Key-words: Control systems, Linearization, Topological equivalence, Grobman-Hartman Theorem

Manuscript submitted to *Journal of differential Equations*, June, 2003.

* `Laurent.Baratchart@sophia.inria.fr`
† Dept. of Mathematics, University of Hawaii, 2565 McCarthy Mall, Honolulu, HI 96822-2273, USA. Email: `mchyba@math.hawaii.edu`.
‡ `Jean-Baptiste.Pomet@sophia.inria.fr`

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Sur les équivalents du théorème de Grobmann-Hartmann pour les systèmes de contrôle

Résumé : On s'intéresse au problème de la linéarisation topologique des systèmes de contrôle, c'est-à-dire l'équivalence locale à un système linéaire commandable via des transformations bi-continues qui ne sont pas forcément différentiables. On prouve d'une part que, s'agissant de transformations ponctuelles, la linéarisation topologique implique la linéarisation différentiable, au moins en dehors de certaines singularités. D'autre part, si l'on autorise les transformations à dépendre du contrôle de manière fonctionnelle, on arrive a prouver un équivalent du théorème de Grobman-Hartman.

Mots-clés : Systèmes de contrôle, linéarisation locale, équivalence topologique, Théorème de Grobman-Hartman

Contents

1 Introduction

In the theory of dynamical systems, it is well-known that, except in some "degenerate" cases, the flow of a nonlinear differential equation near an equilibrium point "looks like" the flow of its tangent approximation. More precisely, the Grobman-Hartman theorem says that these flows are conjugate around a hyperbolic equilibrium *via* a local homeomorphism, which is however not differentiable in general. Consequently it is impossible, *locally around an equilibrium*, to distinguish a hyperbolic nonlinear system from a linear one on the basis of *qualitative* phenomena.

For control systems, it has been known for some time [12, 9] that smooth linearizability, namely local equivalence to a linear control system by means of a *diffeomorphic* change of variables (also known as *static feedback linearization*), requires very restrictive conditions. But in view of the Grobman-Hartman theorem, it is natural to ask whether there may exist, in the neighborhood of an equilibrium, a topological although perhaps non-differentiable transformation that conjugates the input-to-state behavior of a “generic” system to that of a linear controllable system. If yes, this would mean that a non-linear control system, under appropriate non-degeneracy conditions, cannot be distinguished qualitatively from its linear approximation, at least locally around an equilibrium point. Also, it would imply that systems with linear dynamics, whose nonlinear character lies in their input-to-state and state-to-output functions only, are generic models for control systems in the neighborhood of an equilibrium.

The first objective of the present paper is to provide an essentially negative answer to this question, by showing (*cf* Theorem 6.2) that a smooth control system which is locally topologically linearizable at some point (that may or may not be an equilibrium) is in fact locally “quasi-smoothly” linearizable. The definition of this notion is technical, *cf* Definition 5.6; it coincides with smooth linearizability away from singularities, while, at singular points, deciding whether it implies the existence of a linearizing homeomorphism which is smooth (but not necessarily a diffeomorphism) remains for the authors an open issue that raises an intriguing question in differential topology (*cf* section 5.4). The proof of Theorem 6.2 relies on classical results concerning orbits of families of smooth vector fields, first given in [24], that we recall and slightly expand in Appendix D. Incidentally, the essential use that we make of results from [24] is the most compelling reason why we assume that the control systems under consideration are of class C^∞ and not merely C^k.

The second objective of the paper is to derive, along the same lines as the Grobman-Hartman theorem for (non-controlled) differential equations, some positive results on the local linearization of control systems of class C^1 with respect to the state variable. These do not contradict the above mentioned "negative" results because the notion of conjugacy is here much weaker: either the control itself is generated by a finite dimensional dynamical system, or else the transformations depend both on the past and on the future values of the control using an abstract representation of the system as a flow on some functional space in the style of [5]. These two results are derived from a common abstract principle (Theorem 7.1, the second main result of the paper) stating that, when controls are generated by a flow (*i.e.* a one parameter group of homeomorphism on some topological space), then, under quantitative hyperbolicity assumptions, it can be linearized *via* transformations that are continuously parameterized by the elements of this topological space.

Although they deal with closely related matters, the two parts devoted respectively to these two objectives are essentially *independent* from each other, except for some common background and motivations in section 2 and a few general definitions in section 3.1. The first part spans sections 3.1 to 6, and the second part is contained in section 7.

The organization of the paper is as follows. Section 2 briefly recalls classical facts on the local linearization of ordinary differential equations. Section 3 defines conjugation of *control systems* (under a homeomorphism, diffeomorphism, etc...) and states some basic properties of conjugating maps. Section 4 reviews (topological, smooth, linear) conjugacy between *linear* control systems after [2, 27]. Section 5 recalls the known conditions for smooth conjugacy of a smooth nonlinear control system to a linear one, and introduces the notion of quasi-smooth linearization. Section 6 states the first main result of the paper, namely local topological linearizability implies local quasi-smooth linearizability for smooth control systems. The proof consists of two lemmas gathered in appendix F, that in turn depend on section 3, on results from [24] that are stated in appendix D, and on technical lemmas from Appendices A and E. Section 7 states the second main result: Theorem 7.1, whose proof requires results from Appendices B, and C, and that implies two Grobman-Hartman theorems of weak type, stated in subsections 7.2 and 7.3.

2 Differentiable and topological linearization for ordinary differential equations

The question of linearization of a differential equation by a smooth change of coordinates around an equilibrium point is a very old one. At the beginning of the twentieth century, H. Poincaré already identified the obstructions to the existence of a *formal* change of coordinates that removes all the nonlinear terms. These obstructions are the so-called resonances, see *e.g.* [8]. They are, of course, obstructions to smooth linearization as well. It turns out [19], if no eigenvalue of the Jacobian is purely imaginary, that the absence of resonances is also sufficient for smooth (but not real analytic) linearization.

Next, if one allows conjugation *via* a topological but not necessarily differentiable homeomorphism, the Grobman-Hartman theorem shows that every ordinary differential equation with no purely imaginary eigenvalue of the Jacobian can be locally linearized, *i.e.* resonances are no longer an obstruction. Our point of departure will be a brief review of this classical result after fixing some notation. Consider the differential equation

$$\dot{x}(t) = f(x(t)), \tag{2.1}$$

where $f \in C^1(U, I\!R^n)$ and U is an open subset of $I\!R^n$. Assume that $x_0 \in U$ is an equilibrium, i.e. $f(x_0) = 0$. The linearized system associated to (2.1) near x_0 is

$$\dot{x}(t) = Ax(t) - Ax_0 \tag{2.2}$$

where $A = Df(x_0)$ is the derivative of f at x_0. The equilibrium x_0 is said to be *hyperbolic* if the matrix A has no purely imaginary eigenvalue. Systems (2.1) and (2.2) are called *topologically conjugate* at x_0 if there exist neighborhoods V, W of x_0 in U and a homeomorphism

$h : V \to W$ mapping the trajectories of (2.1) in V onto the trajectories of (2.2) in W in a time-preserving manner : for each $x \in V$, we should have

$$h \circ \phi_t(x) = e^{At}\big(h(x) - h(x_0)\big) + h(x_0)$$

provided that $\phi_\rho(x) \in V$ for $0 \le \rho \le t$, where ϕ_t denotes the flow of (2.1). The Grobman-Hartman theorem now goes as follows [8]:

Theorem 2.1 (Grobman-Hartman) *Under the assumption that x_0 is an hyperbolic equilibrium point,* (2.1) *is topologically conjugate to* (2.2) *at x_0.*

This theorem entails that the only invariant under local topological conjugacy around a hyperbolic equilibrium is the number of eigenvalues with positive real part in the Jacobian matrix, counting multiplicity. Indeed, it is well-known (*cf* [1]) that the linear system $\dot{x} = Ax$ where A has no pure imaginary eigenvalue is topologically conjugate to the linear system $\dot{x} = DX$ where D is diagonal with diagonal entries ± 1, the number of occurrences of $+1$ being the number of eigenvalues of A with positive real part, counting multiplicity.

3 Preliminaries on topological equivalence for control systems

3.1 Definitions

Consider now two *control systems* :

$$\dot{x} = f(x,u)\ ,\quad x \in I\!R^n\ ,\quad u \in I\!R^m\ , \tag{3.1}$$
$$\dot{z} = g(z,v)\ ,\quad z \in I\!R^{n'}\ ,\quad v \in I\!R^{m'}\ , \tag{3.2}$$

or expanded in coordinates :

$$\begin{array}{rcl} \dot{x}_1 &=& f_1(x_1,\ldots,x_n,u_1,\ldots,u_m) \\ \vdots && \\ \dot{x}_n &=& f_n(x_1,\ldots,x_n,u_1,\ldots,u_m) \end{array} \qquad \begin{array}{rcl} \dot{z}_1 &=& g_1(z_1,\ldots,z_{n'},v_1,\ldots,v_{m'}) \\ \vdots && \\ \dot{z}_{n'} &=& g_{n'}(z_1,\ldots,z_{n'},v_1,\ldots,v_{m'}) \end{array}$$

where x (resp. z) is called the *state* and u (resp. v) the *control.* Note that it is not mandatory for f and g to *actually depend* on u or v: if, for instance, f does not depend on u, then equation (3.1) reduces to an ordinary differential equation which is often called a *non-controlled system* in this setting in order to stress its independence from the control.

Henceforth we assume that f and g are at least continuous, in order to guarantee the existence of trajectories for control systems like (3.1) and (3.2), but *any additional regularity assumption will be stated explicitly.* We hasten to say that the results of sections 5 and 6 are obtained under the hypothesis that f and g are smooth (i.e. infinitely differentiable), while those in section 7 require them to be continuously differentiable with respect to the first variable. However, in the present section devoted to general background on topological equivalence, we shall try to keep assumptions to a minimum. We always suppose that f and g are defined on the whole of $I\!R^n \times I\!R^m$ and $I\!R^{n'} \times I\!R^{m'}$ respectively, but this is no loss of generality to us because almost every result that we prove (with one exception pointed out in Remark 7.13) is local with respect to x, u, z, and v, so that f and g can be extended using partitions of unity outside some neighborhood of the arguments under consideration without affecting the conclusions. This extra elbow-room is used at several places to avoid dealing with non-complete vector fields.

Definition 3.1 *By* a solution of (3.1) that remains in an open set $\Omega \subset I\!R^{n+m}$, *we mean a mapping γ defined on a real interval I, say*

$$\begin{array}{rccccl} \gamma : & I & \rightarrow & \Omega & & \\ & t & \mapsto & \gamma(t) & = & (\,\gamma_{\mathrm{I}}(t)\,,\,\gamma_{\mathrm{II}}(t)\,) \end{array} \tag{3.3}$$

with $\gamma_{\mathrm{I}}(t) \in I\!R^n$ and $\gamma_{\mathrm{II}}(t) \in I\!R^m$, such that :

- *γ is measurable, locally bounded, and γ_{I} is absolutely continuous,*
- *whenever $[T_1, T_2] \subset I$, we have :*

$$\gamma_{\mathrm{I}}(T_2) - \gamma_{\mathrm{I}}(T_1) \;=\; \int_{T_1}^{T_2} f(\,\gamma_{\mathrm{I}}(t)\,,\,\gamma_{\mathrm{II}}(t)\,)\,\mathrm{d}t\;. \tag{3.3b}$$

Solutions of (3.2) that remain in $\Omega' \subset I\!R^{n'+m'}$ are likewise defined to be mappings

$$\begin{array}{rccccl} \gamma' : & I & \rightarrow & \Omega' & & \\ & t & \mapsto & \gamma'(t) & = & (\,\gamma'_{\mathrm{I}}(t)\,,\,\gamma'_{\mathrm{II}}(t)\,) \end{array} \tag{3.4}$$

having the corresponding properties with respect to g.

Remark 3.2 *Observe that Definition 3.1 assigns a definite value to $\gamma_{\mathrm{II}}(t)$ for* each $t \in I$. *Of course, since γ_{I} remains a solution to* (3.3b) *when the control γ_{II} gets redefined over a set of measure 0, one could identify two control functions whose values agree a.e. on I, as is customary in integration theory. However, these values are in any case subject to the constraint that $\gamma(t) \in \Omega$ for* every $t \in I$, *and altogether we find it more convenient to adopt Definition 3.1.*

If $(\bar{x}, \bar{u})$ is a point in Ω and $\mathcal{U}$ a neighborhood of $\bar{u}$ such that $\bar{x} \times \mathcal{U} \subset \Omega$, any measurable and locally bounded map $\gamma_{\mathrm{II}} : J \rightarrow \mathcal{U}$, where J is a real interval, gives rise on a possibly smaller interval $I \subset J$ to a solution γ of (3.1) that remains in Ω in the sense of Definition 3.1, subject to the initial condition $\gamma_{\mathrm{I}}(0) = \bar{x}$. This is due to the continuity of f and the local boundedness of γ_{II} [4, Ch. 2, Theorem 1.1]. This solution however needs not be unique, even locally around $(\bar{x}, \bar{u})$, as is well-known of the Cauchy problem for continuous vector fields. One condition that ensures uniqueness of the solution on some maximal interval of definition, once γ_{II} and $\bar{x}$ are fixed, is that f be *locally Lipschitz in the first argument* on Ω, in other words that each $(\bar{x}, \bar{u}) \in \Omega$ has a neighborhood $\mathcal{N}$ such that $\|f(x', u) - f(x, u)\| \leq c\|x' - x\|$ for some constant c whenever (x, u) and (x', u) lie in $\mathcal{N}$, see *e.g.* [22, Theorem 54]. The local Lipschitz condition is of course automatically satisfied if f is smooth.

In the terminology of control, a solution in the sense of Definition 3.1 would be termed *open loop* to emphasize that the value of the control at time t is a function of time only, namely that $\gamma_{\mathrm{II}}(t)$ bears no relation to the state x whatsoever. A central concept in control theory, though, is that of *closed loop* or *feedback* control, where the value of the control at time t is computed from the corresponding value of the state, namely is of the form $\alpha(x(t))$. To make a formal definition of a feedback defined on an arbitrary open set, we need one more piece of notation : if $\Omega \subset I\!\!R^n \times I\!\!R^m$ is open, we let $\Omega_{I\!\!R^n} \subset I\!\!R^n$ denote the projection of Ω onto the first factor, so that Ω becomes a fibered space over $\Omega_{I\!\!R^n}$ using the natural projection $\pi_n : \Omega \to \Omega_{I\!\!R^n}$ that selects the first n components.

Definition 3.3 *Given an open set $\Omega \subset I\!\!R^{n+m}$, a* feedback *on Ω is a continuous mapping $\alpha : \Omega_{I\!\!R^n} \to I\!\!R^m$ such that $(x, \alpha(x)) \in \Omega$ for all $x \in \Omega_{I\!\!R^n}$. A* smooth *feedback on Ω is one which is a smooth mapping.*

Naturally associated to the control system (3.1) and to the feedback α is the continuous vector field f_α on $\Omega_{I\!\!R^n}$ defined by

$$f_\alpha(x) \;=\; f(x, \alpha(x)) \;. \tag{3.5}$$

A feedback is nothing but a mapping α such that $x \mapsto (x, \alpha(x))$ is a continuous section of the natural fibration $\pi_n : \Omega \to \Omega_{I\!\!R^n}$. Of course, there are sets Ω whose topology prevents the existence of any feedback. However, if there is one there are plenty, among which smooth feedbacks are uniformly dense. This is the content of the next proposition, that will be used in the proof of Theorem 6.2. To fix notations, let us agree throughout that the symbol $\| \ \|$ designates the Euclidean norm on $I\!\!R^\ell$ irrespectively of the positive integer ℓ, while $B(x, r)$ stands for the open ball centered at x of radius r and $\overline{B}(x, r)$ for the corresponding closed ball.

Proposition 3.4 *Let Ω be open in $I\!\!R^{n+m}$, and $\alpha : \Omega_{I\!\!R^n} \to I\!\!R^m$ be a feedback on Ω. To each $\varepsilon > 0$, there is a smooth feedback $\beta : \Omega_{I\!\!R^n} \to I\!\!R^m$ such that $\|\alpha(x) - \beta(x)\| < \varepsilon$ for $x \in \Omega_{I\!\!R^n}$.*

Proof. Let $\varnothing = \mathcal{K}_0 \subset \mathcal{K}_1 \subset \mathcal{K}_2 \ldots \subset \mathcal{K}_k \subset \mathcal{K}_{k+1} \ldots$ be an increasing sequence of compact subsets of $\Omega_{I\!\!R^n}$, each of which contains the previous one in its interior, whose union is all of $\Omega_{I\!\!R^n}$. For each $x \in \Omega_{I\!\!R^n}$, define the integer

$$k(x) \;\stackrel{\Delta}{=}\; \min\{k;\ x \in \mathcal{K}_k\} \;. \tag{3.6}$$

To each k, by the continuity of α and the compactness of $\mathcal{K}_k$, there is $\mu_k > 0$ such that

$$x \in \mathcal{K}_k \;\Rightarrow \left\{ \begin{array}{l} \bullet\ B(x, \mu_k) \times \mathrm{Conv}\left\{ \alpha\big(B(x, \mu_k)\big) \right\} \;\subset\; \Omega\,, \\ \bullet\ \forall u_1, u_2 \in \mathrm{Conv}\left\{ \alpha\big(B(x, \mu_k)\big) \right\},\ \|u_1 - u_2\| < \varepsilon\,, \end{array} \right. \tag{3.7}$$

where the symbol Conv designates the convex hull. In addition, we may assume that the sequence (μ_k) is non increasing.

Denote by $\overset{\circ}{\mathcal{K}}_k$ the interior of $\mathcal{K}_k$, set $\mathcal{D}_k = \mathcal{K}_k \setminus \overset{\circ}{\mathcal{K}}_{k-1}$ for $k \geq 1$, and cover the compact set $\mathcal{D}_k$ with a finite collection $\mathcal{B}_k$ of open balls having the following properties :

- each of these balls is centered at a point of $\mathcal{D}_k$ and is contained in the open set $\overset{\circ}{\mathcal{K}}_{k+1} \setminus \mathcal{K}_{k-2}$ (with the convention that $\mathcal{K}_{-1} = \varnothing$),
- each of these balls has radius at most $\dfrac{\mu_{k+1}}{2}$.

The union $\mathcal{B} = \bigcup_{k\geq 1} \mathcal{B}_k$ is a countable locally finite collection of open balls that covers $\Omega_{\mathbb{R}^n}$, and it has the property that *every ball in $\mathcal{B}$ is included in $B(x, \mu_{k(x)})$ as soon as it contains x*. Let B_j, for $j \in \mathbb{N}$, enumerate $\mathcal{B}$, and h_j be a smooth partition of unity where h_j has support $\mathrm{supp} h_j \subset B_j$. If we pick $x_j \in B_j$ for each j, the map $\beta : \Omega_{\mathbb{R}^n} \to \mathbb{R}^m$ defined by

$$\beta(x) = \sum_{j\in\mathbb{N}} h_j(x)\alpha(x_j) \tag{3.8}$$

is certainly smooth. In addition, since by construction x_j belongs to $B(x, \mu_{k(x)})$ whenever $h_j(x) \neq 0$, we get that $\beta(x)$ lies in the convex hull of $\alpha\big(B(x,r)\big)$ for some $r < \mu_{k(x)}$, and therefore, from (3.7) and (3.6), that $(x, \beta(x)) \in \Omega$ and $\|\alpha(x) - \beta(x)\| < \varepsilon$. Hence β is a smooth feedback on Ω such that $\|\alpha(x) - \beta(x)\| < \varepsilon$ for all $x \in \Omega_{\mathbb{R}^n}$. $\square$

We now turn to the notion of conjugacy for control systems, which is the central topic of the paper.

Definition 3.5 *Let*

$$\begin{array}{rcccl} \chi : & \Omega & \to & \Omega' & \\ & (x,u) & \mapsto & \chi(x,u) & = \ (\, \chi_{\mathrm{I}}(x,u)\,,\, \chi_{\mathrm{II}}(x,u)\,) \end{array} \tag{3.9}$$

be a bijective mapping between two open subsets of $\mathbb{R}^{n+m}$ and $\mathbb{R}^{n'+m'}$ respectively. We say that χ conjugates *systems* (3.1) *and* (3.2) *if, for any real interval I, a map $\gamma : I \to \Omega$ is a solution of* (3.1) *that remains in Ω if, and only if, $\chi \circ \gamma$ is a solution of* (3.2) *that remains in Ω'.*

We say that systems (3.1) *and* (3.2) *are* topologically conjugate *over the pair Ω,Ω' if there exists a homeomorphism $\chi : \Omega \to \Omega'$ that conjugates the two systems; we say that these are* smoothly conjugate *if, in addition, χ and χ^{-1} are smooth.*

We say that system (3.1) *is* locally topologically (resp. smoothly) conjugate to system (3.2) at[1] $(\bar{x}, \bar{u}) \in \mathbb{R}^{n+m}$ *if the two systems are topologically (resp. smoothly) conjugate over a pair Ω, Ω', where Ω is a neighborhood of $(\bar{x}, \bar{u})$.*

[1] It would be more natural to say that system (3.1) **at $(\bar{x}, \bar{u})$**$\in \mathbb{R}^{n+m}$ is locally conjugate to system (3.2) **at $(\bar{x}', \bar{u}')$**$\in \mathbb{R}^{n+m}$ if the two systems are conjugate over a pair Ω, Ω', where Ω is a neighborhood of $(\bar{x}, \bar{u})$ and Ω' is a neighborhood of $(\bar{x}', \bar{u}')$.

However, the present definition is correct, and prescribing $(\bar{x}', \bar{u}')$ would increase complexity and add no information here.

In case there is no control (*i. e.* $m = m' = 0$) so that neither u nor χ_{II} appear in equation (3.9), Definition 3.5 coincides with the classical notion of local topological conjugacy for ordinary differential equations, and will serve as a formal definition in this case too.

Remark 3.6 *The definition is invariant under linear time re-parameterization, namely :* if $\chi : \Omega \to \Omega'$ conjugates systems (3.1) and (3.2), then for any $\lambda \in I\!R$ (if $\lambda < 0$, this reverses time), the map χ also conjugates the systems

$$\dot{x} = \lambda f(x,u) \quad \text{and} \quad \dot{z} = \lambda g(z,v) \ .$$

Indeed, this is trivial for $\lambda = 0$, otherwise, if $t \mapsto (x(t), u(t))$ is a solution of $\dot{x} = \lambda f(x,u)$ on a time-interval $[t_1, t_2]$, and $\tilde{x}(t)$ and $\tilde{u}(t)$ denote respectively $x(t/\lambda)$ and $u(t/\lambda)$, then $t \mapsto (\tilde{x}(t), \tilde{u}(t))$ is a solution of (3.1) *on $[\lambda t_1, \lambda t_2]$, hence χ sends $(\tilde{x}(t), \tilde{u}(t))$ to $(\tilde{z}(t), \tilde{v}(t))$ satisfying $\dot{\tilde{z}}(t) = g(\tilde{z}(t), \tilde{v}(t))$. Consequently, χ maps $(x(t), u(t))$ to $(z(t), v(t)) = ((\tilde{z}(\lambda t), \tilde{v}(\lambda t)))$, which is a solution of $\dot{z} = \lambda g(z,v)$.*

3.2 Alternative definitions

In the literature, there seems to be no general agreement on what should be called a solution of a control system, nor on the concept of equivalence. Let us quickly mention some of the notions in use.

The work [27], devoted to the topological classification of *linear* control systems, is reviewed in section 4.2. It makes for another definition of solutions that privileges the state x with respect to the control u, namely a solution there consists merely of $x(t)$ instead of $(x(t), u(t))$. This amounts to replacing Definition 3.1 by one where a solution is a map γ_{I} *such that there exists* a map γ_{II} for which γ is a solution in the sense of that definition. This new set of solutions is the projection on the x factor of our set of solutions, and if two different $u(t)$ produce the same $x(t)$ then Definition 3.1 distinguishes between two solutions whereas this alternative point of view would see only one. It is natural in this case to define a corresponding notion of conjugacy, similar to the one in Definition 3.5, where the conjugating map $z = \phi(x)$ conjugates only the state. To avoid confusion in the rare instances where we use this alternative notion, we shall speak then of *x-solution* and *x-conjugacy* respectively.

In [5], a control system like (3.1) is viewed as a flow on the product space $I\!R^n \times \mathcal{U}$, where $\mathcal{U}$ is a functional space of admissible controls whose dynamics is induced by the time-shift. See section 7.3 for more precisions and a Grobman-Hartman theorem in this setting.

In [3], control systems are defined abstractly without reference to differential equations, as maps $(x(0), u(.)) \mapsto x(.)$ that satisfy certain axioms. In that paper a notion of topological equivalence is proposed that amounts, for systems like (3.1), to the existence of a certain triangular map on the product $I\!R^n \times \mathcal{U}$, see Remark 7.14 in section 7.3.

Until section 7, we will only deal with transformations on *finite dimensional* point sets; we therefore postpone any further discussion on functional approaches until that section. As far as finite dimensional point sets transformations are concerned, since, for nonlinear control systems, local results have to be stated locally with respect both to x and u, we favor Definition 3.5 of conjugacy, and x-conjugacy we will no longer be mentioned x-conjugacy, except in Remark 3.10 where we will stress the link between the notions of *conjugacy* and *x-conjugacy*, after we have established some general properties of conjugating maps.

3.3 Some properties of conjugating maps

When (3.1) and (3.2) are topologically conjugate over some pair of open sets, it follows immediately from Brouwer's invariance of the domain (see *e.g.* [16]) that $n' + m' = n + m$. Proposition 3.7 below asserts that more in fact is true, namely conjugating homeomorphisms must have a triangular structure, which implies in particular that both $n' = n$ and $m' = m$.

Proposition 3.7 *With the notations of Definition 3.5, suppose that* (3.1) *and* (3.2) *are topologically conjugate via a homeomorphism* $\chi : \Omega \to \Omega'$. *Then* $n = n'$, $m = m'$, *and* χ_{I} *depends only on* x:

$$\chi(x, u) \quad = \quad (\, \chi_{\mathrm{I}}(x) \, , \, \chi_{\mathrm{II}}(x, u) \,) \ . \tag{3.10}$$

Moreover, $\chi_{\mathrm{I}} : \Omega_{\mathbb{R}^n} \to \Omega'_{\mathbb{R}^n}$ *is a homeomorphism. Here, one should recall the notation* $\Omega_{\mathbb{R}^n}$ *that was introduced before Definition 3.3.*

Proof. Let $\bar{x}$, $\bar{u}$, $\bar{u}'$ be such that $(\bar{x}, \bar{u})$ and $(\bar{x}, \bar{u}')$ belong to Ω. Let further $x(t)$ be a solution[2] to (3.1) with $x(0) = \bar{x}$ and

$$\begin{aligned} u(t) &= \bar{u} \ \text{ if } \ t \le 0, \\ u(t) &= \bar{u}' \ \text{ if } \ t > 0 \ . \end{aligned}$$

By conjugacy, $z(t) = \chi_{\mathrm{I}}(x(t), u(t))$ is a solution to (3.2) with v given by $v(t) = \chi_{\mathrm{II}}(x(t), u(t))$, for $t \in (-\epsilon, \epsilon)$ and some $\epsilon > 0$. In particular $\chi_{\mathrm{I}}(x(t), u(t))$ is continuous in t so its values at 0^+ and 0^- are equal. Hence $\chi_{\mathrm{I}}(\bar{x}, \bar{u}) = \chi_{\mathrm{I}}(\bar{x}, \bar{u}')$ so that $\chi_{\mathrm{I}} : \Omega_{\mathbb{R}^n} \to \Omega'_{\mathbb{R}^{n'}}$ is well defined and continuous. Similarly, $\left(\chi^{-1}\right)_{\mathrm{I}}$ induces a continuous inverse $\Omega'_{\mathbb{R}^{n'}} \to \Omega_{\mathbb{R}^n}$. $\square$

In view of Proposition 3.7, we will only consider conjugacy between systems having the same number of states and inputs. Hence the distinction between (n, m) and (n', m') from now on disappears.

Remark 3.8 *Taking into account the triangular structure of* χ *in Proposition 3.11, one may describe conjugation as resulting from a change of coordinates in the state-space (upon setting* $z = \chi_{\mathrm{I}}(x)$*) and then feeding the system with a function both of the state and of a new control variable* v *(upon setting* $u = (\chi^{-1})_{\mathrm{II}}(z, v)$*), in such a way that the correspondence* $(x, u) \mapsto (z, v)$ *is invertible. In the language of control, this is known as a* static feedback transformation, *and two systems conjugate in the sense of Definition 3.5 would be termed* equivalent under static feedback. *This notion has received considerable attention (see for*

[2] This solution is not necessarily unique since here f and g are merely assumed to be continuous.

instance [11]), albeit only in the differentiable *case (i.e. when the triangular transformation χ is a diffeomorphism).*

Note that when χ and χ^{-1} are smooth, one may replace Definition 3.5, which is based on the notion of solutions, *by a differential formula translating the way in which χ transforms the* equations*: systems* (3.1) *and* (3.2) *are locally smoothly conjugate on some domain if, and only if,*

$$g(\chi(x,u)) \quad = \quad \frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\, f(x,u) \tag{3.11}$$

holds on this domain.

In the proof of Proposition 3.7, we only used conjugacy on a very small class of solutions, namely those corresponding to piecewise constant controls with a single discontinuity. This raises the question whether smaller classes of solutions than prescribed in Definition 3.1 are still sufficiently rich to check for conjugacy. Under mild conditions on f and g, as we will see in the forthcoming proposition, conjugacy essentially holds if it is granted for a class of inputs that locally uniformly approximates piecewise continuous functions, and this fact will be of technical use in the proof of Lemma F.2 which is itself an ingredient of the proof of Theorem 6.2. Let us fix some terminology here, and agree that a function $I \to \mathbb{R}^m$, where I is a real interval, is called *piecewise continuous* if it is continuous except possibly at *finitely many* interior points of I where it has limits from both sides and is either right or left continuous. If in addition the function is constant (resp. affine, smooth) on every open interval not containing a discontinuity point, we say that it is *piecewise constant* (resp. *piecewise affine, smooth*).

Proposition 3.9 *Assume that f and g are continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and locally Lipschitz-continuous with respect to their first argument. Let $\chi : \Omega \to \Omega'$ be a homeomorphism between two open subsets of $\mathbb{R}^{n+m}$, and denote by Ω_{II} and Ω'_{II} respectively the open subsets of $\mathbb{R}^m$ obtained by projecting Ω and Ω' onto the second factor. Let further $\mathcal{C}$ and $\mathcal{C}'$ be collections of locally bounded measurable functions $\mathbb{R} \to \mathbb{R}^m$ whose restrictions $\mathcal{C}|_J$ and $\mathcal{C}'|_J$ to any compact interval J contain in their respective closures, for the topology of uniform convergence, the set of all piecewise continuous functions $J \to \Omega_{\mathrm{II}}$ and $J \to \Omega'_{\mathrm{II}}$ respectively. If χ maps every solution* (3.3) *of* (3.1) *such that $\gamma_{\mathrm{II}}(t) \in \mathcal{C}|_I$ to a solution of* (3.2) *while, conversely, χ^{-1} maps every solution* (3.4) *of* (3.2) *such that $\gamma'_{\mathrm{II}}(t) \in \mathcal{C}'|_I$ to a solution of* (3.1)*, then the restriction of χ to any relatively compact open subset $O \subset \Omega$ conjugates systems* (3.1) *and* (3.2) *over the pair O, $\chi(O)$.*

It may come as a disappointment that Proposition 3.9 only asserts conjugacy to hold over relatively compact open subsets of Ω, but it is unclear to the authors whether systems (3.1) and (3.2) should be conjugate on the whole of Ω without additional regularity assumptions on the latter.

Proof. Let us first show that

$$\left.\begin{array}{l}\text{for any solution } \gamma : I \to \Omega \text{ of (3.1) } \textit{such that } \gamma_{\mathrm{II}} \textit{ is} \\ \textit{piecewise continuous}, \chi \circ \gamma \text{ is a solution to (3.2).}\end{array}\right\} \tag{3.12}$$

Since the property of being a solution is local with respect to time, we may suppose that I is a compact interval. Then, there is an open set $\mathcal{O}$ and a compact set $\mathcal{K}$ such that $\gamma(I) \subset \mathcal{O} \subset \mathcal{K} \subset \Omega$. By the hypothesis on $\mathcal{C}$, there exists a sequence of functions $\gamma_{\mathrm{II},k} : I \to \mathbb{R}^m$ converging uniformly to γ_{II} such that $\gamma_{\mathrm{II},k} \in \mathcal{C}|_I$. Define for each $k \in \mathbb{N}$ a time-varying vector field X^k by $X^k(t,x) = f(x, \gamma_{\mathrm{II},k}(t))$. By the continuity of f, this sequence converges uniformly on compact subsets of $I \times \mathbb{R}^n$ to $X(t,x) = f(x, \gamma_{\mathrm{II}}(t))$; moreover, since γ_{II} is bounded (being piecewise continuous) $\gamma_{\mathrm{II},k}$ is also bounded, thus the local Lipschitz character of $f(x,u)$ with respect to x implies by compactness that $X(t,x)$ and $X^k(t,x)$ are themselves locally Lipschitz with respect to x on $I \times \mathcal{O}_{\mathbb{R}^n}$. Pick $t_0 \in I$ and apply Lemma A.3 with $I = [t_1, t_2]$, $x_0 = \gamma_{\mathrm{I}}(t_0)$, and $\mathcal{U} = \mathcal{O}_{\mathbb{R}^n}$. This yields, say for $k > K$, that the solution $\gamma_{\mathrm{I},k}$ to the Cauchy problem

$$\dot{\gamma}_{\mathrm{I},k}(t) = X^k(t, \gamma_{\mathrm{I},k}(t)) \qquad \gamma_{\mathrm{I},k}(t_0) = \gamma_{\mathrm{I}}(t_0)$$

maps I into $\mathcal{O}_{\mathbb{R}^n}$ and that the sequence $(\gamma_{\mathrm{I},k})_{k>K}$ converges uniformly on I to γ_{I}. Hence, if we let

$$\gamma_k(t) = (\gamma_{\mathrm{I},k}(t), \gamma_{\mathrm{II},k}(t)),$$

the sequence $(\gamma_k)_{k>K}$ converges to γ, uniformly on I. In particular $\gamma_k(I) \subset \mathcal{K} \subset \Omega$ for k large enough.

Now, since $\gamma_k : I \to \Omega$ is a solution to (3.1) with $\gamma_{\mathrm{II},k} \in \mathcal{C}|_I$, it follows from the hypothesis that $\chi \circ \gamma_k$ is a solution to (3.2) that remains in Ω', *i.e.* with the notations of (3.9) we have, for k large enough,

$$\chi_{\mathrm{I}} \circ \gamma_k(t) \;-\; \chi_{\mathrm{I}} \circ \gamma_k(t_0) \;=\; \int_{t_0}^{t} g(\chi \circ \gamma_k(s))\, \mathrm{d}s, \qquad t \in I. \tag{3.13}$$

By the continuity of χ, the convergence of $\gamma_k(t)$ to $\gamma(t)$, and the fact that g remains bounded on the compact set $\chi(\mathcal{K})$, we can apply the dominated convergence theorem to the right hand-side of (3.13) to obtain in the limit, as $k \to \infty$, that

$$\chi_{\mathrm{I}} \circ \gamma(t) \;-\; \chi_{\mathrm{I}} \circ \gamma(t_0) \;=\; \int_{t_0}^{t} g(\chi \circ \gamma(s))\, \mathrm{d}s, \qquad t \in I.$$

Thus $\chi \circ \gamma : I \to \mathbb{R}^{n+m}$ is a solution to (3.2) that remains in Ω', thereby proving (3.12).

The next step is to observe from (3.12) that, since piecewise constant controls are in particular piecewise continuous, the proof of Proposition 3.7 applies to show that $\chi : \Omega \to \Omega'$ has a triangular structure of the form (3.10).

With (3.12) and (3.10) at our disposal, let us now prove the proposition in its generality. Choose an arbitrary open subset $\mathcal{O}$ with compact closure $\overline{\mathcal{O}}$ in Ω, and fix two compact subsets $\mathcal{K}$ and $\mathcal{K}_1$ of Ω such that

$$\mathcal{O} \;\subset\; \overline{\mathcal{O}} \;\subset \overset{\circ}{\mathcal{K}} \subset\; \mathcal{K} \;\subset \overset{\circ}{\mathcal{K}}_1 \subset\; \mathcal{K}_1 \;\subset\; \Omega.$$

where $\overset{\circ}{\mathcal{K}}$ stands for the *interior* of $\mathcal{K}$.

Let $\gamma : I \to \mathcal{O}$ be a solution of (3.1). We need to prove that $\chi \circ \gamma$ is a solution to (3.2) and again, since the property of being a solution is local with respect to time, we may suppose that I is compact. Notations being as in (3.3), it follows by definition of a solution that γ_{II} is a bounded measurable function $I \to I\!R^m$. We shall proceed as before in that we again approximate γ by a sequence γ_k of trajectories of (3.1) that are mapped by χ to trajectories of (3.2). This time, however, the approximation process is slightly more delicate, because it is no longer granted by the hypothesis on $\mathcal{C}$ but it will rather depend on general point-wise approximation properties to measurable functions by continuous ones.

By the compactness of $\mathcal{K}$, there is $\varepsilon_{\mathcal{K}} > 0$ such that

$$(x,u) \in \mathcal{K} \ \Rightarrow\ B\big((x,u)\,,\,\varepsilon_{\mathcal{K}}\big) \subset \overset{\circ}{\mathcal{K}}_1\ . \tag{3.14}$$

Let $u_{\gamma_{\mathrm{I}}} : I \to I\!R^m$ be an auxiliary function with the following properties :

(i) $u_{\gamma_{\mathrm{I}}}$ is piecewise constant on I,

(ii) $(\xi(t), u_{\gamma_{\mathrm{I}}}(t)) \in \overset{\circ}{\mathcal{K}}_1$ for all $t \in I$ and every map $\xi : I \to I\!R^n$ that satisfies

$$\sup_{t\in I} \|\xi(t) - \gamma_{\mathrm{I}}(t)\| < \varepsilon_{\mathcal{K}}/2. \tag{3.15}$$

Such a function $u_{\gamma_{\mathrm{I}}}$ certainly exists. Indeed, by definition of a solution, γ_{I} is absolutely continuous thus *a fortiori* continuous $I \to I\!R^n$, and therefore we know for each $t \in I$ that the set

$$\gamma_{\mathrm{I}}^{-1}\big(B(\gamma_{\mathrm{I}}(t), \varepsilon_{\mathcal{K}}/2)\big)$$

is an open neighborhood of t in I, hence a disjoint union of open intervals in I one of which contains t; call this particular interval U_t. By the compactness of I, we may cover the latter with finitely many intervals U_{t_j} for $1 \le j \le \nu$. Let now $j(t)$ denote, for each $t \in I$, the smallest index $j \in \{1, \dots, \nu\}$ such that $t \in U_{t_j}$. Then, the map

$$u_{\gamma_{\mathrm{I}}}(t) = \gamma_{\mathrm{II}}(t_{j(t)})$$

clearly satisfies (i), and since $(\gamma_{\mathrm{I}}(t_{j(t)}), \gamma_{\mathrm{II}}(t_{j(t)})) \in \mathcal{O} \subset \mathcal{K}$, it follows from (3.14) and the fact that $\|\gamma_{\mathrm{I}}(t) - \gamma_{\mathrm{I}}(t_{j(t)})\| < \varepsilon_{\mathcal{K}}/2$ by definition of $j(t)$ that $u_{\gamma_{\mathrm{I}}}$ also satisfies (ii).

Next, recall that γ_{II} is a bounded measurable function $I \to I\!R^m$ so, by Lusin's theorem [21, Theorem 2.23] applied component-wise, there is, for every integer $k \ge 1$, a continuous function $h_k : I \to I\!R^m$ that coincides with γ_{II} outside some set $\mathcal{T}_k \subset I$ of Lebesgue measure strictly less than $1/k^2$, and in addition such that

$$\sup_{t\in I} \|h_k(t)\| \le \sqrt{m} \sup_{t\in I} \|\gamma_{\mathrm{II}}(t)\|. \tag{3.16}$$

Put

$$E_k = \{t \in I;\ (\gamma_{\mathrm{I}}(t), h_k(t)) \notin \overset{\circ}{\mathcal{K}}\}.$$

Since h_k is continuous E_k is compact, and since $\gamma(I) \subset \mathcal{O} \subset \overset{\circ}{\mathcal{K}}$ it is clear that $E_k \subset \mathcal{T}_k$ hence E_k has Lebesgue measure strictly less than $1/k^2$. Consequently, by the outer regularity of Lebesgue measure, E_k can be covered by finitely many open real intervals $I_{k,1}, \ldots, I_{k,N_k}$ whose lengths add up to no more than $1/k^2$.

We now define the sequence of functions $\gamma_{\mathrm{II},k}$ on I by setting, for $k \geq 1$,

$$\begin{aligned} \gamma_{\mathrm{II},k}(t) &= h_k(t) \text{ if } t \in I \setminus \textstyle\bigcup_{j=1}^{N_k} I_{k,j}, \\ \gamma_{\mathrm{II},k}(t) &= u_{\gamma_{\mathrm{I}}}(t) \text{ if } t \in \textstyle\bigcup_{j=1}^{N_k} I_{k,j}. \end{aligned} \tag{3.17}$$

By construction $\gamma_{\mathrm{II},k}$ is piecewise continuous, and uniformly bounded independently of k in view of (3.16) and the fact that $u_{\gamma_{\mathrm{I}}}$, being piecewise constant, is bounded. Moreover, as $\sum_{k\geq 1} 1/k^2 < \infty$, the measure of the set $\cup_{j=1}^{N_k} I_{k,j}$ is the general term, indexed by k, of a convergent series, hence almost every $t \in I$ belongs at most to finitely many of these sets so that $\gamma_{\mathrm{II},k}$ converges point-wise a.e. to γ_{II} on I as $k \to \infty$.

Redefine now $X^k(t,x) = f(x, \gamma_{\mathrm{II},k}(t))$, $X(t,x) = f(x, \gamma_{\mathrm{II}}(t))$, and observe from what we just said and the continuity of f that $X^k(t,x)$ converges to $X(t,x)$ when $k \to \infty$, locally uniformly with respect to $x \in \mathcal{O}_{I\!R^n}$, as soon as $t \notin E$ where $E \subset I$ is a set of zero measure which is independent of k. Moreover, again from the boundedness of $\gamma_{\mathrm{II},k}$, γ_{II} and the local Lipschitz character of f, we have that $X^k(t,x)$, $X(t,x)$ are locally Lipschitz with respect to x. Pick $t_0 \in I$ and apply Lemma A.3 with $\mathcal{U} = \mathcal{O}_{I\!R^n}$, $I = [t_1, t_2]$, and $x_0 = \gamma_{\mathrm{I}}(t_0)$. We get, say for $k > K$, that the solution $\gamma_{\mathrm{I},k}$ to the Cauchy problem

$$\dot{\gamma}_{\mathrm{I},k}(t) = X^k(t, \gamma_{\mathrm{I},k}(t)), \qquad \gamma_{\mathrm{I},k}(t_0) = \gamma_{\mathrm{I}}(t_0)$$

is defined over I, maps the latter into $\mathcal{O}_{I\!R^n}$, and that the sequence $(\gamma_{\mathrm{I},k})_{k>K}$ converges uniformly on $[t_1, t_2]$ to γ_{I}.

We claim that $\gamma_k(t) = (\gamma_{\mathrm{I},k}(t), \gamma_{\mathrm{II},k}(t))$ lies in $\overset{\circ}{\mathcal{K}}_1$ for all $t \in I$ when k is so large that

$$\sup_{t \in I} \|\gamma_{\mathrm{I},k}(t) - \gamma_{\mathrm{I}}(t)\| < \varepsilon_{\mathcal{K}}/2. \tag{3.18}$$

Indeed, if $t \in \cup_j I_{k,j}$, this follows automatically from definition (3.17) by property (ii) of $u_{\gamma_{\mathrm{I}}}$; if $t \notin \cup_j I_{k,j}$, then $(\gamma_{\mathrm{I}}(t), h_k(t)) \in \overset{\circ}{\mathcal{K}}$ by the very definition of $\cup_j I_{k,j}$, and since $\gamma_k(t) = (\gamma_{\mathrm{I},k}(t), h_k(t))$ in this case, we deduce from (3.14) and (3.18) that $\gamma_k(t) \in \overset{\circ}{\mathcal{K}}_1$. *This proves the claim.*

Altogether, we have shown that $\gamma_k : I \to \overset{\circ}{\mathcal{K}}_1$ is a solution of (3.1) as soon as k is large enough, with $\gamma_{\mathrm{II},k}$ a piecewise continuous function on I by construction. By (3.12), we now deduce that, for k large enough, $\gamma'_k = \chi \circ \gamma_k$ is a solution of (3.2) that stays in Ω'. Let us block-decompose γ'_k into

$$\gamma'_{\mathrm{I},k}(t) \ = \ \chi_{\mathrm{I}}(\, \gamma_{\mathrm{I},k}(t)\,)\ , \quad \gamma'_{\mathrm{II},k}(t) \ = \ \chi_{\mathrm{II}}(\, \gamma_{\mathrm{I},k}(t)\, , \, \gamma_{\mathrm{II},k}(t)\,)\ ,$$

where we have taken into account the triangular structure of χ. That $\gamma'_k : I \to \Omega'$ is a solution of (3.2) means exactly that

$$\gamma'_{\mathrm{I},k}(t) \; - \; \gamma'_{\mathrm{I},k}(t_0) \;\; = \;\; \int_{t_0}^{t} g(\, \gamma'_{\mathrm{I},k}(s) \, , \, \gamma'_{\mathrm{II},k}(s) \,) \mathrm{d}s, \qquad t \in I. \tag{3.19}$$

Due to the continuity of χ, the functions $\gamma'_{\mathrm{I},k}$ and $\gamma'_{\mathrm{II},k}$ respectively converge uniformly and point-wise almost everywhere to $\gamma'_{\mathrm{I}} = \chi_{\mathrm{I}} \circ \gamma_{\mathrm{I}}$ and $\gamma'_{\mathrm{II}} = \chi_{\mathrm{II}} \circ \gamma$ on I. Since g is bounded on the compact set $\chi(\mathcal{K}_1)$ that contains $\gamma_k(I)$ for k large enough, we get on the one hand, by dominated convergence, that the right-hand side of (3.19) converges, as $k \to \infty$, to $\int_{t_0}^{t} g(\, \gamma'_{\mathrm{I}}(s) \, , \, \gamma'_{\mathrm{II}}(s) \,)\mathrm{d}s$, and on the other hand that the left-hand side converges to $\gamma'_{\mathrm{I}}(t) - \gamma'_{\mathrm{I}}(t_0)$. Therefore $(\gamma'_{\mathrm{I}}, \gamma'_{\mathrm{II}}) = \chi \circ \gamma : I \to \Omega'$ is a solution of (3.2).

This way we have shown that χ maps any solution of (3.1) that stays in a relatively compact open subset $\mathcal{O}$ of Ω to a solution of (3.2) that stays in Ω'. This achieves the proof, for the converse is obtained symmetrically upon swapping f and g, $\mathcal{C}$ and $\mathcal{C}'$, and replacing χ by χ^{-1}. $\square$

Remark 3.10 *In Section 3.2, we introduced the notion of x-conjugacy, which is somewhat different from the conjugacy used in the present paper. The two propositions established in this section so far will now allow for us to further compare these two notions.*

To begin with, conjugacy implies x-conjugacy by Proposition 3.7. The converse is not true in general, in particular because conjugacy preserves the number of states and inputs by the proposition just quoted, whereas x-conjugacy preserves the dimension of the state (by invariance of domain), but may not preserve the dimension of the input: any control system $\dot{x} = f(x, u)$ with state $x \in I\!R^n$ and input $u \in I\!R^m$ is x-conjugate to the control system $\dot{z} = f(z, \pi_m(v))$ with state $z \in I\!R^n$ and input $v \in I\!R^{m+k}$ where, we recall, π_m denotes the projection onto the first m components. It is therefore necessary to assume *that $n = n'$ and $m = m'$ in order to compare conjugacy and x-conjugacy in a nontrivial way. Under this assumption, x-conjugacy via a* continuously differentiable *homeomorphism ϕ implies conjugacy* locally around points where f and g are Lipschitz-continuous with respect to their first argument and where $\frac{\partial f}{\partial u}$ and $\frac{\partial g}{\partial v}$ exist and have constant rank. *This follows easily from the inverse function theorem, thanks to Proposition 3.9 that shows it is enough to check local conjugacy on piecewise smooth inputs. It is unclear how much these conditions can be relaxed, but the following example shows that conjugacy does not imply x-conjugacy if the constant rank assumption fails, even though $m = m'$ and f, g, ϕ are smooth.*

Indeed, consider the two systems $\dot{x} = u^3 - xu$ and $\dot{z} = v$ with x, u, z, v scalars, i.e. $m = n = m' = n' = 1$. Any absolutely continuous $x(t)$ or $z(t)$ is a x-solution of the corresponding system : for an associated control, take $v(t)$ to be $\dot{z}(t)$ and $u(t)$ to be for instance the smallest root of $u(t)^3 - x(t)u(t) - \dot{x}(t) = 0$. Hence the set of x-solutions is the same for these two systems and they are x-conjugate, via $z = x$. However they are not locally topologically conjugate at $(0,0)$. Indeed, using Proposition 3.7, suppose that a transformation $(z, v) = (\chi_{\mathrm{I}}(x), \chi_{\mathrm{II}}(x, u))$ conjugates these two systems on a neighborhood of $(0,0)$. To any strictly positive x, there are three distinct values $u_1(x)$, $u_2(x)$, $u_3(x)$ for which

$u_i(x)^3 - xu_i(x) = 0$, and these go to zero with x. Consequently, the three constant-maps $t \mapsto (x, u_i(x))$ for $1 \leq i \leq 3$ are three trajectories of the first system whose image will be contained in an arbitrary small neighborhood of $(0,0)$ if x is small enough. Since χ is one-to-one, the constant maps $t \mapsto (\chi_{\mathrm{I}}(x), \chi_{\mathrm{II}}(x, u_i(x)))$ should be three distinct trajectories of the second system for strictly positive but sufficiently small x, which is clearly impossible.

The triangular structure of conjugating homeomorphisms asserted by Proposition 3.7 is to the effect that any such homeomorphism $\chi : \Omega \to \Omega'$ is a fiber preserving map from the bundle $\Omega \to \Omega_{\mathbb{R}^n}$ to the bundle $\Omega' \to \Omega'_{\mathbb{R}^n}$. Since feedbacks are naturally associated to sections of these bundles by Definition 3.3, χ gives rise to a natural transformation from feedbacks on Ω to feedbacks on Ω'. This transformation will prove important enough to deserve a notation : to any feedback α on Ω, we associate a feedback $\chi \blacksquare \alpha$ on Ω' by the formula

$$\chi \blacksquare \alpha(z) \ \triangleq \ \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), \alpha(\chi_{\mathrm{I}}^{-1}(z)))\,. \tag{3.20}$$

We leave it to the reader to check that the properties of an action are satisfied, and in particular that

$$\chi^{-1} \blacksquare (\chi \blacksquare \alpha) \ = \ \alpha\,. \tag{3.21}$$

If the homeomorphism χ in (3.10) conjugates system (3.1) to system (3.2), then it is clear that χ_{I} maps the solutions of the ordinary differential equation $\dot{x} = f_\alpha(x)$ (where the vector field f_α was defined in (3.5)) to the solutions of the ordinary differential equation $\dot{z} = g_{\chi \blacksquare \alpha}(z)$. Indeed if $x(t)$ is a solution of the former, then $(x(t), \alpha(x(t)))$ is a solution of the control system (3.1) in the sense of Definition 3.1 so the conjugacy assumption implies that $(\chi_{\mathrm{I}}(x(t)), \chi_{\mathrm{II}}(x(t), \alpha(x(t))))$ is a solution of (3.2), and setting $z(t) = \chi_{\mathrm{I}}(x(t))$ one clearly has $\chi_{\mathrm{II}}(x(t), \alpha(x(t)))) = \chi \blacksquare \alpha(z(t))$; hence $z(t)$ is a solution to $\dot{z} = g_{\chi \blacksquare \alpha}(z)$ because $(z(t), \chi \blacksquare \alpha(z(t)))$ is a solution of (3.2).

Now, if α_1 and α_2 are two feedbacks on Ω, and the two vector fields f_{α_1} and f_{α_2} are defined on $\Omega_{\mathbb{R}^n}$ by (3.5), we denote their difference by $\delta f_{\alpha_1,\alpha_2}$:

$$\delta f_{\alpha_1,\alpha_2} \ = \ f_{\alpha_1} - f_{\alpha_2}\,. \tag{3.22}$$

Such vector fields are similar to the difference vector fields used in [14], except that we consider arbitrary feedbacks instead of constant ones. These vector fields will play an essential role : the following proposition states that a homeomorphism that conjugates two control systems also conjugates the integral curves of such difference vector fields.

Proposition 3.11 *Suppose that f and g in (3.1) and (3.2) are continuous and locally Lipschitz continuous with respect to their first argument. Assume they are locally topologically conjugate at $(0,0)$ via a homeomorphism $\chi : \Omega \to \Omega'$. Then, notations for χ_{I} and χ_{II} being as in Proposition 3.7, we have for every pair of feedbacks α_1, α_2 on Ω that χ_{I} conjugates any solution of*

$$\dot{x} \ = \ \delta f_{\alpha_1,\alpha_2}(x) \tag{3.23}$$

that remains in $\Omega_{I\!R^n}$ to a solution of

$$\dot{z} \;=\; \delta g_{\chi\blacksquare\alpha_1,\chi\blacksquare\alpha_2}(z) \tag{3.24}$$

that remains in $\Omega'_{I\!R^n}$.

It is perhaps worth emphasizing that the solutions of (3.23) and (3.24) need not be unique since α is merely assumed to be continuous. The following proof carries over to arbitrary linear combinations with constant coefficients of the f_α's, but we only need (in Section 6) the result for differences.

Proof. Let $\eta : [t_1, t_2] \to \Omega_{I\!R^n}$ be an integral curve of $\delta f_{\alpha_1,\alpha_2}$, and set

$$u_1(t) \;=\; \alpha_1(\eta(t)) \;\;, \qquad u_2(t) \;=\; \alpha_2(\eta(t)) \;. \tag{3.25}$$

Let further $\widehat{f} : I\!R^{n+m} \to I\!R^n$ be bounded, continuous and Lipschitz continuous with respect to its first argument, and coincide with f on some compact neighborhood of

$$\eta([t_1,t_2]) \times \Big(\alpha_1(\eta([t_1,t_2])) \bigcup \alpha_2(\eta([t_1,t_2]))\Big) \;.$$

Such a $\widehat{f}$ is easily obtained upon multiplying f by a smooth function with compact support. For $\ell \in I\!N$, let η^ℓ be the solution to the Cauchy problem

$$\eta^\ell(t) \;=\; \eta(t_1) \;+\; \int_{t_1}^{t} G_\ell(\tau, \eta^\ell(\tau))\mathrm{d}\tau \;, \tag{3.26}$$

with

$$\begin{array}{rcl} G_\ell(t,x) &=& 2\,\widehat{f}(x,u_1(t)) \\ && \quad \text{if } t \in [t_1 + \frac{j}{\ell}(t_2-t_1), t_1 + (\frac{j}{\ell}+\frac{1}{2\ell})(t_2-t_1)), \\ G_\ell(t,x) &=& -2\,\widehat{f}(x,u_2(t)) \\ && \quad \text{if } t \in [t_1 + (\frac{j}{\ell}+\frac{1}{2\ell})(t_2-t_1), t_1 + \frac{j+1}{\ell}(t_2-t_1)), \\ G_\ell(t_2,x) &=& -2\,\widehat{f}(x,u_2(t_2)), \qquad 0 \le j \le \ell-1. \end{array} \tag{3.27}$$

The definition of η^ℓ is valid because, since $G_\ell(t,x)$ is bounded and locally Lipschitz with respect to the variable x, the solution to (3.26) uniquely exists.

From Lemma E.1 applied to the case where $X^{1,\ell}(t,x) = \widehat{f}(x,u_1(t))$ and $X^{2,\ell}(t,x) = \widehat{f}(x,u_2(t))$ are in fact independent of ℓ, any accumulation point of the sequence (η^ℓ), say η^∞, is a solution to

$$\dot{\eta}^\infty(t) \;=\; \widehat{f}(\eta^\infty(t),u_1(t)) \;-\; \widehat{f}(\eta^\infty(t),u_2(t)) \;\;, \qquad \eta^\infty(t_1) \;=\; \eta(t_1) \;.$$

Since $\widehat{f}$ is locally Lipschitz continuous with respect to its first argument, the solution to this Cauchy problem is unique and, since f and $\widehat{f}$ coincide at all points $(\eta(t),u_1(t))$ and

$(\eta(t), u_2(t))$, this entails $\eta^\infty = \eta$. Thus (η^ℓ) converges uniformly to η on $[t_1, t_2]$ and, for ℓ large enough, η^ℓ remains a solution of (3.26) if $\widehat{f}$ is replaced by f in (3.27). Moreover, $\eta^\ell([t_1, t_2]) \subset \Omega_{\mathbb{R}^n}$ for ℓ large since the same is true of η. Since χ conjugates the two systems, hence also by Remark 3.6 the systems where f and g are multiplied by 2 or -2, the map $\chi_{\mathrm{I}} \circ \eta^\ell : [t_1, t_2] \to \Omega'_{\mathbb{R}^n}$ is, for ℓ large enough, a solution to

$$\chi_{\mathrm{I}} \circ \eta^\ell(t) \;=\; \chi_{\mathrm{I}} \circ \eta(t_1) \;+\; \int_{t_1}^{t} \widetilde{G}_\ell(\tau, \chi_{\mathrm{I}} \circ \eta^\ell(\tau))\mathrm{d}\tau \tag{3.28}$$

with

$$\begin{aligned}
\widetilde{G}_\ell(t,z) &= 2\,g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), u_1(t))) \\
&\qquad\qquad \text{if } t \in [t_1 + \tfrac{j}{\ell}(t_2 - t_1), t_1 + (\tfrac{j}{\ell} + \tfrac{1}{2\ell})(t_2 - t_1)), \\
\widetilde{G}_\ell(t,z) &= -2\,g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), u_2(t))) \\
&\qquad\qquad \text{if } t \in [t_1 + (\tfrac{j}{\ell} + \tfrac{1}{2\ell})(t_2 - t_1), t_1 + \tfrac{j+1}{\ell}(t_2 - t_1)), \\
\widetilde{G}_\ell(t_2,z) &= -2\,g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), u_2(t_2))).
\end{aligned} \tag{3.29}$$

Since $(\chi_{\mathrm{I}} \circ \eta^\ell)$ converges uniformly to $\chi_{\mathrm{I}} \circ \eta$ by the continuity of χ, replacing g by a bounded and continuous $\widehat{g} : \mathbb{R}^{n+m} \to \mathbb{R}^n$ that coincides with g on a compact neighborhood of

$$\chi_{\mathrm{I}} \circ \eta([t_1, t_2]) \times \Big(\chi_{\mathrm{II}}(\eta([t_1,t_2]), \alpha_1(\eta([t_1,t_2]))) \bigcup \chi_{\mathrm{II}}(\eta([t_1,t_2]), \alpha_2(\eta([t_1,t_2]))) \Big) ,$$

does not affect the validity of (3.28)-(3.29) for ℓ large enough. Lemma E.1 now implies that all accumulation points of the sequence $(\chi_{\mathrm{I}} \circ \eta^\ell)$ in the uniform topology on $[t_1, t_2]$ are solutions of

$$\dot{z} = g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), u_1(t))) \;-\; g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), u_2(t))).$$

Because $\chi_{\mathrm{I}} \circ \eta$ is such an accumulation point, it is by (3.25) a solution to

$$\dot{z} \;=\; g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), \alpha_1(\chi_{\mathrm{I}}^{-1}(z))) \;-\; g(z, \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z), \alpha_2(\chi_{\mathrm{I}}^{-1}(z))) ,$$

which is nothing but (3.24). □

4 The case of linear control systems

4.1 Kronecker indices

A *linear* control systems is a special instance of (3.1), of the form

$$\dot{x} \;=\; Ax \;+\; Bu \tag{4.1}$$

where A and B are constant matrices of dimension $n \times n$ and $n \times m$ respectively. When dealing with linear systems, it is natural to consider an equivalence relation similar to that of Definition 3.5, but where χ is restricted to be a linear isomorphism :

Definition 4.1 *Two linear systems*

$$\dot{x} \;=\; Ax \;+\; Bu \qquad \text{and} \qquad \dot{z} \;=\; \widetilde{A}z \;+\; \widetilde{B}v$$

are linearly conjugate *if and only if any of the following two equivalent properties is satisfied :*

1. *There is a nonempty open set $\Omega \subset \mathbb{R}^{n+m}$, and a linear isomorphism χ of $\mathbb{R}^{n+m}$ whose restriction $\Omega \to \chi(\Omega)$ conjugates the two systems in the sense of Definition 3.5.*
2. *There exist matrices $P \in \mathbb{R}^{n\times n}$, $Q \in \mathbb{R}^{m\times m}$ and $K \in \mathbb{R}^{n\times m}$, with P and Q invertible, such that*

$$\begin{array}{rcl} \widetilde{A} &=& P(A-BK)P^{-1}\;, \\ \widetilde{B} &=& PBQ^{-1}\;. \end{array} \tag{4.2}$$

Since, by Proposition 3.7, a linear conjugating homeomorphism is necessarily of the form $(x,u) \mapsto (Px, Kx+Qu)$, the equivalence between properties 1 and 2 follows at once from differentiating the solutions. Provided it exists, Ω plays absolutely no role in this context since (4.2) implies that the two systems are in fact linearly conjugate on all of $\mathbb{R}^{n+m}$.

Linear conjugacy actually defines an equivalence relation on linear control systems or equivalently on pairs (A,B), for which (4.2) can be read as "(A,B) is equivalent to $(\widetilde{A},\widetilde{B})$". The classification of linear systems under this equivalence relation is well-known [2], and goes as follows. Each equivalence class contains a pair (A_c, B_c) of the form (block matrices) :

$$A_c \;=\; \begin{pmatrix} A_0^c & 0 & \cdots & 0 \\ 0 & A_1^c & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_m^c \end{pmatrix} \;,\quad B_c \;=\; \begin{pmatrix} 0 & \cdots & 0 \\ b_1^c & \ddots & \vdots \\ 0 & \ddots & 0 \\ \vdots & 0 & b_m^c \end{pmatrix} \tag{4.3}$$

where

$$A_i^c \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & & & 0 \\ & & & & \ddots & 1 \\ 0 & \cdots & & & \cdots & 0 \end{pmatrix}_{(\kappa_i\times\kappa_i)} \;,\qquad b_i^c \;=\; \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}_{(\kappa_i\times 1)} ,\quad 1\leq i\leq m. \tag{4.4}$$

The integers $(\kappa_1,\ldots,\kappa_m)$ are called the controllability indices of the control system, also known as the Kronecker indices of the matrix pencil (A,B), while A_0^c is a square matrix of dimension $n-(\kappa_1+\cdots+\kappa_m)$ that may be assumed in Jordan canonical form. Note that $\kappa_1+\cdots+\kappa_m \leq n$, and if $\kappa_1+\cdots+\kappa_m = n$ there is no A_0^c ; also, it may well happen that $\kappa_i = 0$, in which case A_i^c and b_i^c are empty and do not occur in (4.3) to the effect that there are less than m blocks beyond A_0^c. Normalizing so that

$$\kappa_1 \geq \cdots \geq \kappa_m \geq 0,$$

and ordering the Jordan blocks in some way, there is one and only one such normal form per equivalence class. A complete set of invariants is then the list of Kronecker indices and the spectral invariants of the matrix A_0^c.

With the natural partition $z = (Z_0, Z_1, \ldots, Z_m)$ corresponding to the block decomposition (4.3), the control system associated to the pair (A_c, B_c) reads

$$\dot{Z}_0 \;=\; A_0 Z_0\;,\quad \dot{Z}_1 \;=\; A_1 Z_1 \;+\; u_1 b_1^c\;,\quad \cdots\quad ,\quad \dot{Z}_m \;=\; A_m Z_m \;+\; u_m b_m^c\;,$$

where Z_0 is missing if $\kappa_1 + \cdots + \kappa_m = n$ and Z_i is missing if $\kappa_i = 0$. Because it is not influenced at all by the controls, Z_0 is sometimes called the non-controllable part of the state. In this paper, we are only interested in controllable linear systems, namely :

Definition 4.2 *A linear control system* (4.1) *is said to be* controllable *if, and only if, the following two equivalent properties are satisfied :*

1. *There is no bloc A_0^c in the associated normal form* (4.3).
2. *Kalman's criterion for controllability :*

$$\mathrm{Rank}(B, AB, \ldots, A^{n-1}B) = n.$$

To see the equivalence of the two properties, observe that the $n - \kappa_1 - \cdots - \kappa_m$ first rows of the matrix P that puts (A, B) into canonical form (i.e. $z = Px$) form a basis of the smallest dual subspace that annihilates the columns of B and at the same time is invariant under right multiplication by A, i.e. they are a basis of the left kernel of $(B, AB, \ldots, A^{n-1}B)$. For controllable linear systems, the only invariant under linear conjugacy is thus the ordered list of Kronecker indices. These can be computed from $(B, AB, \ldots, A^{n-1}B)$ as follows : if we put

$$\begin{array}{rcll} r_j &=& \mathrm{Rank}(B, AB, \ldots, A^{j-1}B)\;, & r_0 = 0\;,\\ s_j &=& r_j - r_{j-1}\;, & s_0 = m\;, \end{array} \tag{4.5}$$

then s_j does not increase with j and a moment's thinking will convince the reader that the number of Kronecker indices that are equal to i is $s_i - s_{i+1}$, or equivalently that κ_k is the number of s_j's that are $\geq k$.

To us, it will be more convenient not to use the above normal form, but rather the following permutation of it. Let ρ be the smallest integer such that $s_\rho = 0$, so that

$$0 \;=\; s_\rho \;<\; s_{\rho-1} \;\leq\; s_{\rho-2} \;\leq\; \cdots \;\leq\; s_1 \;\leq\; s_0 \;=\; m\;,$$

with $\displaystyle\sum_{j\geq 1} s_j = n$. From these we define, for $0 \leq i \leq \rho$:

$$\sigma_i \;=\; \sum_{j\geq i} s_j, \tag{4.6}$$

so that in particular $\sigma_\rho = 0$, $\sigma_{\rho-1} = s_{\rho-1} > 0$, $\sigma_1 = n$ and $\sigma_0 = n+m$. Note that, from (4.5), $\sigma_i = n - r_{i-1}$ for $i \geq 1$. We shall write our controllable canonical form as $\dot z = A_c z + B_c v$ with

$$A_c = \left(\begin{array}{c|c|c|c|c|c} 0 & J^{s_{\rho-1}}_{s_{\rho-2}} & & & & \\ \hline & 0 & J^{s_{\rho-2}}_{s_{\rho-3}} & & & \\ \hline & & 0 & J^{s_{\rho-3}}_{s_{\rho-4}} & & \\ \hline & & & \ddots & \ddots & \\ \hline & & & & 0 & J^{s_2}_{s_1} \\ \hline & & & & & 0 \end{array}\right), \quad B_c = \left(\begin{array}{c} 0 \\ \hline J^{s_1}_{s_0} \end{array}\right) \tag{4.7}$$

where for any integers r and s with $s \leq r$, J^s_r is the $s \times r$ matrix

$$J^s_r = \left(\begin{array}{c|c} I_s & 0 \end{array}\right) \tag{4.8}$$

where I_s is the $s \times s$ identity matrix.

4.2 Topological classification of linear control systems

In [27], which is devoted to the topological classification of *linear* control systems and uses the notion of x-conjugacy rather than conjugacy (*cf* Section 3.2 and Remark 3.10), the following result is proved:

Theorem 4.3 (Willems [27]) *If two linear control systems $\dot x = Ax + Bu$ and $\dot z = \widetilde{A}z + \widetilde{B}v$ are topologically x-conjugate, then they have the same list of Kronecker indices, and the non-controllable blocks A_0^c and $\widetilde{A}_0^c$ in their respective canonical forms* (4.3) *are such that the two linear differential equations $\dot X_0 = A_0^c X_0$ and $\dot Z_0 = \widetilde{A}_0^c Z_0$ are topologically equivalent.*

As pointed out in Remark 3.10, topological conjugacy implies topological x-conjugacy but not conversely. However, for linear control systems having the same number m of inputs, Theorem 4.3 implies that these notions are equivalent. Indeed, if two systems are respectively brought into their canonical form (4.3) by a linear change of variable on $\mathbb{R}^{n+m}$, and if in addition they are x-conjugate, then their non-controllable parts are topologically equivalent while the remaining blocks are identical by equality of the Kronecker indices. Hence, both in the above theorem and in the corollary below, one may use indifferently "x-conjugate" or "conjugate"

Corollary 4.4 *If two linear systems $\dot{x} = Ax + Bu$ and $\dot{z} = \widetilde{A}z + \widetilde{B}v$ are topologically conjugate and one of them is controllable, then the other one is controllable too and they are linearly conjugate.*

Proof. Controllability is preserved, since Kronecker indices are by the theorem. Linear conjugacy follows, as we saw that the list of Kronecker indices is a complete invariant for controllable systems. □

In the controllable case, the results in section 6 can be viewed, in a local setting, as a generalization of Corollary 4.4 to the case where only one of the two systems is linear, to the effect that they are locally quasi-smoothly conjugate (*cf* Definition 5.6).

5 Differentiable linearization for control systems

In this section and the next one, we consistently assume that f in system (3.1) is smooth, i.e. of class C^∞.

5.1 Differentiable linearization

Definition 5.1 *The system* (3.1) *is said to be* locally smoothly linearizable at $(\bar{x}, \bar{u}) \in \mathbb{R}^{n+m}$ *if it is locally smoothly conjugate, in the sense of Definition 3.5, to a linear controllable system $\dot{z} = Az + Bv$ (see Definition 4.2).*

Proposition 5.2 *Let $(\bar{x}, \bar{u})$ be an equilibrium point of* (3.1), *i.e. $f(\bar{x}, \bar{u}) = 0$, and let $\overline{A} = \partial f/\partial x\,(\bar{x}, \bar{u})$, $\overline{B} = \partial f/\partial u\,(\bar{x}, \bar{u})$ so that :*

$$f(x, u) \;=\; \overline{A}\,(x - \bar{x}) \;+\; \overline{B}\,(u - \bar{u}) \;+\; \varepsilon(x - \bar{x}, u - \bar{u})\;, \tag{5.1}$$

where ε is little $\mathrm{o}(\|\mathrm{x} - \bar{\mathrm{x}}\| + \|\mathrm{u} - \bar{\mathrm{u}}\|)$.

If system (3.1) *is locally smoothly linearizable at $(\bar{x}, \bar{u})$, then*

1. its linear approximation $(\overline{A}, \overline{B})$ is controllable (cf. Definition 4.2), and

2. the system is smoothly conjugate to $(\overline{A}, \overline{B})$ at $(\bar{x}, \bar{u})$.

Proof. Let χ be a local diffeomorphism conjugating system (3.1) to $\dot{z} = Az + Bv$ at $(\bar{x}, \bar{u})$, and observe from (3.11) in Remark 3.8 that smooth linearizability translates into

$$\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\, f(x, u) \quad = \quad A\,\chi_{\mathrm{I}}(x) \;+\; B\,\chi_{\mathrm{II}}(x, u)\;. \tag{5.2}$$

If we write f as in (5.1), and if we set $\overline{P} = \frac{\partial \chi_{\mathrm{I}}}{\partial x}(\bar{x})$, $\overline{K} = \frac{\partial \chi_{\mathrm{II}}}{\partial x}(\bar{x},\bar{u})$, $\overline{Q} = \frac{\partial \chi_{\mathrm{II}}}{\partial u}(\bar{x},\bar{u})$, we get by differentiating (5.2) with respect to x and u at $(\bar{x},\bar{u})$, using the relation $f(\bar{x},\bar{u})=0$, that

$$\overline{P}\,\overline{A} \;=\; A\overline{P} + B\overline{K}\;, \qquad \overline{P}\,\overline{B} \;=\; B\overline{Q}\,.$$

Since $\overline{P}$ and $\overline{Q}$ are square invertible matrices by the triangular structure of χ displayed in (3.10), this implies that the linear systems (A,B) and $(\overline{A},\overline{B})$ are linearly conjugate, see (4.2). Since (A,B) is controllable by definition so is $(\overline{A},\overline{B})$, thereby achieving the proof. □

As pointed out in Remark 3.8, smooth conjugation, also termed equivalence under static feedback, has been substantially studied. Local smooth linearizability as in Definition 5.1 coincides with linearizability by smooth static feedback as described in [10, 18], a characterization of which was given in [12, 9] for systems that are affine with respect to the control, and more generally for smooth systems in [26]. In order to recall this characterization in a slightly different manner, which is more convenient to us, we define for $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and each $u \in \mathbb{R}^m$ a vector field f_u on $\mathbb{R}^m$ by the formula:

$$f_u(x) \;=\; f(x,u), \tag{5.3}$$

and for L any linear map we let $\operatorname{Ran} L$ denote its range. Furthermore, we set

$$\Delta_0(x,u) \;=\; \operatorname{Ran}\frac{\partial f}{\partial u}(x,u). \tag{5.4}$$

Theorem 5.3 ([12, 9, 26]) *When f is of class C^∞, the control system* (3.1) *is locally smoothly linearizable at $(\bar{x},\bar{u})$ if, and only if, there are open neighborhoods $\mathcal{X}$ and $\mathcal{U}$ of $\bar{x}$ and $\bar{u}$ in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively such that the following conditions are satisfied.*

1. *$\Delta_0(x,u)$ does not depend on u for $(x,u) \in \mathcal{X}\times\mathcal{U}$.*
2. *The rank of $\dfrac{\partial f}{\partial u}(x,u)$ is constant in $\mathcal{X}\times\mathcal{U}$.*
3. *Defining on $\mathcal{X}$ the distribution Δ_0 by $\Delta_0(x) = \Delta_0(x,u)$ — this is possible if point 1 holds true — and inductively the flag of distributions (Δ_k) by :*

$$\Delta_{k+1} \;=\; \Delta_k + [\,f_{\bar{u}}\,,\,\Delta_k\,]\,, \tag{5.5}$$

then each Δ_k for $0 \le k \le n-1$ is integrable (i.e. has constant dimension over $\mathbb{R}$ and is closed under Lie bracket) and the rank of Δ_{n-1} is n.

Strictly speaking, condition 3 supersedes 2 for the constancy of dimension is implicit in the integrability condition. Before we come to the proof of Theorem 5.3, in order to prepare the ground for Theorem 5.9 to come, let us reformulate the former in a slightly different form by introducing, for each $(x,u) \in \mathcal{X}\times\mathcal{U}$, a subspace $D(x,u)$ that will later replace

$\operatorname{Ran} \partial f/\partial u\,(x,u)$ when the constant rank condition (2) happens to fail. First, we define the subset $\mathcal{L}_{x,u} \subset \mathbb{R}^n$ (not a vector subspace) by :

$$y \in \mathcal{L}_{x,u} \quad \Leftrightarrow \quad \exists (h_n) \in (\mathbb{R}^m)^{\mathbb{N}}, \quad \lim_{n\to\infty} h_n = 0 \quad \text{and} \quad \lim_{n\to\infty} \frac{f(x,u+h_n)-f(x,u)}{\|f(x,u+h_n)-f(x,u)\|} = y; \tag{5.6}$$

subsequently we put

$$D(x,u) \quad = \quad \operatorname{Span}_{\mathbb{R}} \mathcal{L}_{x,u} \; . \tag{5.7}$$

In words, $D(x,u)$ is the vector space spanned by all limit directions of straight lines through $f(x,u)$ and $f(x,u')$ as u' approaches u in $\mathbb{R}^m$; it is of common use in stratified geometry to generalize the notion of tangent space. Note that the set $\mathcal{L}_{x,u}$ depends on the norm used in (5.6), but the subspace $D(x,u)$ does not.

Proposition 5.4 *If $f : \mathbb{R}^{n+m} \to \mathbb{R}^n$ is of class C^∞, we have that*

$$D(x,u) \quad \supset \quad \operatorname{Ran} \partial f/\partial u\,(x,u) \tag{5.8}$$

and equality holds at every (x,u) where the rank of $\partial f/\partial u\,(x,u)$ is locally constant with respect to u.

Assuming Proposition 5.4 for a while, we get as an immediate corollary a version of Theorem 5.3 in which condition 2 is no longer redundant :

Corollary 5.5 *Theorem 5.3 is still valid if, instead of (5.4), we define*

$$\Delta_0(x,u) \quad = \quad D(x,u) \; . \tag{5.9}$$

Proof of Theorem 5.3. Define on $\mathbb{R}^{n+m}$ the vector field F_0 by

$$F_0 \; : \quad (x,u) \;\mapsto\; \begin{pmatrix} f(x,u) \\ 0 \end{pmatrix} , \tag{5.10}$$

and the distribution $\mathcal{H}_{-1}$, whose first integrals are all the x coordinates :

$$\mathcal{H}_{-1}(x,u) \quad = \quad \operatorname{Span}_{\mathbb{R}} \{ \frac{\partial}{\partial u_1}(x,u), \ldots, \frac{\partial}{\partial u_m}(x,u) \} \tag{5.11}$$

where $\partial/\partial u_1, \ldots, \partial/\partial u_m$ customarily denote the natural coordinate vector fields; define also, for $k \geq 0$, the distribution $\mathcal{H}_k$ by

$$\mathcal{H}_k \quad = \quad \mathcal{H}_{k-1} \; + \; [\, F_0 \, , \, \mathcal{H}_{k-1} \,] \; . \tag{5.12}$$

The characterization of smooth local linearization by static feedback given in [26] is that all the distributions $\mathcal{H}_k$ be integrable in some neighborhood of $(\bar{x}, \bar{u})$, and that the rank of $\mathcal{H}_{n-1}$ be $n+m$.

Now, it is straightforward to check that points 1 and 2 in the theorem are equivalent to $\mathcal{H}_0$ being involutive (*i.e.* closed under Lie bracket) and of constant dimension around $(\bar{x}, \bar{u})$, this dimension being the rank of $\partial f/\partial u$. Moreover, an easy induction then shows that

$$\mathcal{H}_k = \left(\Delta_k \times \{0\}\right) \oplus \mathcal{H}_{-1}, \qquad k \geq 0.$$

Thus the statement of the theorem is indeed equivalent to [26], as desired. □

Proof of Proposition 5.4. The inclusion (5.8) is clear since any element of $\mathrm{Ran}\,\partial f/\partial u(x,u)$ can be written $\partial f/\partial u\,(x,u).h$ with $h \in \mathbb{R}^m$, and, if it is nonzero, one has

$$\frac{\partial f/\partial u\,(x,u).h}{\|\partial f/\partial u\,(x,u).h\|} = \lim_{t\to 0^+} \frac{f(x,u+th)-f(x,u)}{\|f(x,u+th)-f(x,u)\|}\ .$$

To get the reverse inclusion when the rank of $\partial f/\partial u\ (x,u)$ is locally constant, we use the constant rank-theorem to the effect that the partial map $u' \mapsto f(x,u')$ may, in suitable coordinates around u, be factored as $\varphi \circ i_{r,n} \circ \pi_r$ where φ is a local diffeomorphism around $f(x,u)$ in $\mathbb{R}^n$, r is the rank of $\partial f/\partial u\,(x,u)$, the map $\pi_r : \mathbb{R}^m \to \mathbb{R}^r$ is the canonical projection on the first r components and the inclusion $i_{r,n} : \mathbb{R}^r \to \mathbb{R}^n$ maps $(w_1,\ldots,w_r)$ to $(w_1,\ldots,w_r,0,\ldots,0)$. Therefore, up to a local change of coordinates around u in $\mathbb{R}^m$ which clearly leaves $D(x,u)$ and $\mathrm{Ran}\,\partial f/\partial u\,(x,u)$ unchanged, we can assume that $f(x,u')$ depends only on $\pi_r(u')$ and that

$$\|f(x,u')-f(x,u)\| \geq c\|\pi_r(u'-u)\| \tag{5.13}$$

for some constant $c > 0$. Take $y \in D(x,u)$. By definition there is a sequence (h_n) satisfying (5.6). Since $f(x,u+h_n)$ depends only (for x fixed) on $\pi_r(u+h_n)$, we may replace h_n by any element of $\mathbb{R}^m$ having the same r first components : assume that the $m-r$ last components of h_n are zero, so that $\|h_n\| = \|\pi_r(h_n)\|$. Then, we can write

$$y = \lim_{n\to\infty} \frac{f(x,u+h_n)-f(x,u)}{\|h_n\|}\,\frac{\|\pi_r(h_n)\|}{\|f(x,u+h_n)-f(x,u)\|}. \tag{5.14}$$

Let h be a limit point of $h_n/\|h_n\|$ in $\mathbb{R}^n$. By definition of the derivative, we see, up to a subsequence, that the first ratio on the right-hand side of (5.14) converges to $\partial f/\partial u\,(x,u).h$. Moreover the second ratio is bounded from above by virtue of (5.13), and therefore, extracting another subsequence if necessary, we may assume it converges to some real number. It is now apparent that $y \in \mathrm{Ran}\,\partial f/\partial u\,(x,u)$ as desired. □

5.2 Quasi differentiable linearization

In Theorem 6.2 to come, the relevant notion is not exactly smooth linearization but a slight variant thereof that we now introduce.

Definition 5.6 *We say that system* (3.1) *is* locally quasi-smoothly linearizable at $(\bar{x}, \bar{u})$ *if there exists an open neighborhood* Ω *of* $(\bar{x}, \bar{u})$ *in* $\mathbb{R}^{n+m}$ *and a homeomorphism*

$$\begin{array}{rccccl} \chi : & \Omega & \to & \Omega' & & \\ & (x,u) & \mapsto & \chi(x,u) & = & (\, \chi_{\mathrm{I}}(x) \, , \, \chi_{\mathrm{II}}(x,u) \,) \end{array} \tag{5.15}$$

such that

1. χ *conjugates system* (3.1) *to a linear controllable system* $\dot{z} = Az + Bv$*, in the sense of Definition 3.5,*
2. $\chi_{\mathrm{I}} : \Omega_n \to \Omega'_n$ *is a smooth (i.e.* C^∞*) diffeomorphism.*

We say that the system is quasi-smoothly linearizable on Ω *if it is quasi-smoothly linearizable at every point of the latter.*

Smooth and quasi-smooth linearizations do not coincide, as evidenced by the system

$$\dot{x} \;=\; u^3 \qquad u \in \mathbb{R}, \; x \in \mathbb{R}, \tag{5.16}$$

which is not smoothly linearizable at the origin by Proposition 5.2 because its linear approximation is not controllable, but which is quasi-smoothly linearizable at the origin for the homeomorphism $(x,u) \mapsto (z,v) = (x, u^3)$ conjugates (5.16) to $\dot{z} = v$ and, although it is not a diffeomorphism, its first component obviously is $\mathbb{R} \to \mathbb{R}$.

In the above example, the conjugating homeomorphism was not a diffeomorphism, but it was a smooth map. In fact, we have no counter-example to :

Conjecture 5.7 *If* f *is of class* C^∞ *and system* (3.1) *is quasi-smoothly linearizable at* $(\bar{x}, \bar{u}) \in \mathbb{R}^{n+m}$*, then there exist a neighborhood* Ω *of* $(\bar{x}, \bar{u})$ *together with a linearizing homeomorphism* $\chi : \Omega \to \Omega'$ *satisfying Definition 5.6 and having the additional property that* ***it is a smooth map****.*

Again, the conjecture with "smooth map" replaced by "smooth diffeomorphism" would be defeated by example (5.16). We are going to show that this conjecture is true at "generic" points, but in general we have no definite opinion about it. See section 5.4 for an open question, stated in purely differential-toological terms, which is equivalent to this conjecture. To further understand quasi-smooth linearization, let us define at each point $(x,u) \in \mathbb{R}^{n+m}$ two integers $\nu(x,u)$ and $\widetilde{\nu}(x,u)$ by the formulas :

$$\nu(x,u) = \mathrm{Rank}\, \frac{\partial f}{\partial u}(x,u) \;, \qquad \widetilde{\nu}(x,u) = \inf_{O \in \mathcal{N}(x,u)} \; \sup_{(x',u') \in O} \nu(x',u') \;, \tag{5.17}$$

where $\mathcal{N}(x,u)$ stands for the set of all open neighborhoods of (x,u) in $\mathbb{R}^{n+m}$. As maps $\Omega \to \mathbb{N}$, these have the following properties : $\widetilde{\nu} \geq \nu$, ν is lower semi-continuous, $\widetilde{\nu}$ is upper semi continuous, and both are locally constant on an open dense set.

Proposition 5.8 *If the smooth system* (3.1) *is quasi-smoothly linearizable at* $(\bar{x}, \bar{u}) \in \Omega$, *then the homeomorphism* $\chi : \Omega \to \Omega'$ *and the matrices* A, B *in Definition 5.6 must satisfy :*

1. *the map* $B\chi_{\text{II}} : \Omega \to I\!R^m$ *is of class* C^∞,
2. $\widetilde{\nu}$ *is constant on* Ω, *specifically* $\widetilde{\nu}(x, u) = \text{Rank}B$ *for all* $(x, u) \in \Omega$.

It follows from the proposition that conjecture 5.7 holds true if we suppose in addition that $\widetilde{\nu}(\bar{x}, \bar{u}) = m$, because χ itself is smooth in this case. Indeed χ_I is smooth by definition of quasi-smooth linearization, while χ_{II} is also smooth since $B\chi_{\text{II}}$ is smooth and B has full column rank. In particular, conjecture 5.7 holds true for real analytic f if $n \geq m$ and $\partial f / \partial u$ has rank m in at least one point.

Proof of Proposition 5.8 Computing $\dot{z}$ at the origin of a trajectory starting from $(x, u) \in \Omega$, we get by the smoothness of χ_{I} that (5.2) holds for all $(x, u) \in \Omega$. This makes it clear that $B\chi_{\text{II}}$ is smooth which proves (1). To establish (2), it is enough in view of (5.17) to check the following two facts :

$$\begin{array}{l} \nu(x,u) \leq \text{Rank}B \ \text{ for all } (x,u) \text{ in } \Omega\,, \\ \text{the set } \{(x,u) \in \Omega,\ \nu(x,u) = \text{Rank}B\} \text{ is dense in } \Omega\,. \end{array} \tag{5.18}$$

Now, since $B\chi_{\text{II}}$ is smooth, differentiating (5.2) with respect to u yields

$$\frac{\partial \chi_{\text{I}}}{\partial x}(x)\frac{\partial f}{\partial u}(x,u) \quad = \quad \frac{\partial (B\chi_{\text{II}})}{\partial u}(x,u)\ , \tag{5.19}$$

which clearly implies the first half of (5.18). The second half we prove by contradiction, supposing that there is an open subset $\mathcal{O} \subset \Omega$ on which ν is no larger than $s - 1$ where $s = \text{Rank}B$. Let ρ be the maximum value that ν assumes on $\mathcal{O}$, so that ρ is an integer strictly smaller than s. By lower semi-continuity ν is locally constant equal to ρ on some open set $\mathcal{V} \subset \mathcal{O}$, and from (5.19) we see that $B\chi_{\text{II}}$ is smooth of constant rank ρ on $\mathcal{V}$. But since χ is open the map $(x, u) \mapsto (\chi_{\text{I}}(x), B\chi_{\text{II}}(x, u))$ maps $\mathcal{V}$ onto an open subset of $\text{Ran}I_n \times B$ which is a $n + s$-dimensional linear manifold, and at the same time this map is smooth on with derivative of rank $n + \rho < n + s$ on $\mathcal{V}$, thereby contradicting the constant rank theorem. □

We turn to a characterization of quasi-smooth linearizability that runs parallel to that of smooth linearizability given in Theorem 5.3-Corollary 5.5.

Theorem 5.9 *Suppose that* f *in* (3.1) *is of class* C^∞. *The system* (3.1) *is locally quasi-smoothly linearizable at* $(\bar{x}, \bar{u})$ *if, and only if, there are open neighborhoods* $\mathcal{X}$ *and* $\mathcal{U}$ *of* $\bar{x}$ *and* $\bar{u}$ *in* $I\!R^n$ *and* $I\!R^m$ *respectively such that conditions (1) and (3) of Theorem 5.3 are met with* Δ_0 *given by* $\Delta_0(x, u) = D(x, u)$ *(as in* (5.9), *Corollary 5.5), and, instead of condition 2, it holds that*

(2′) The smooth mapping

$$F \;:\; \begin{array}{ccc} \mathcal{X}\times\mathcal{U} & \to & \mathcal{X}\times I\!R^n \\ (x,u) & \mapsto & (\,x\,,\,f(x,u)\,) \end{array} \tag{5.20}$$

restricts to a C^0 *fibration*[3] $\mathcal{W}\to F(\mathcal{W})$ *with fiber* $I\!R^{m-\widetilde{\nu}(\bar{x},\bar{u})}$ *on some neighborhood* $\mathcal{W}$ *of* $(\bar{x},\bar{u})$ *in* $\mathcal{X}\times\mathcal{U}$.

We postpone the proof for a short while to make a few observations. Firstly, comparing Theorem 5.9 and Corollary 5.5 obviously yields :

Corollary 5.10 *Suppose that a smooth system* (3.1) *is locally quasi-smoothly linearizable at* $(\bar{x},\bar{u})$*; then it is locally smoothly linearizable at* $(\bar{x},\bar{u})$ *if, and only if, the rank of* $\partial f/\partial u$ *is constant around* $(\bar{x},\bar{u})$.

In particular, we deduce the following :

Corollary 5.11 *If a smooth system* (3.1) *is locally quasi-smoothly linearizable at* $(\bar{x},\bar{u})$*, then it is locally smoothly linearizable at points that form an open dense subset of a neighborhood of* $(\bar{x},\bar{u})$.

Secondly, it is worth noting the following :

Remark 5.12 *Smooth linearizability at an equilibrium implies, by Proposition 5.2, smooth conjugacy* to the linear approximation. *The situation is different regarding quasi-smooth linearizability. Let us give examples of quasi-smoothly linearizable systems which are* not *conjugate to their linear approximation. This is the case of system* (5.16)*, whose linear approximation at the origin is not controllable whereas the system is nevertheless quasi-smoothly conjugate to a linear controllable system at this point. Apart from such degenerate cases, there also exist systems that are quasi-smoothly linearizable at some point with controllable linear approximation there, and still they are not conjugate to this linear approximation. An example when* $m=n=2$ *is given by :*

$$\dot{x}_1 \;=\; u_1 \;\;,\qquad \dot{x}_2 \;=\; x_1 \,+\, u_2^3 \;,$$

This system is quasi-smoothly conjugate at $(0,0)$ *to*

$$\dot{z}_1 \;=\; v_1 \;\;,\qquad \dot{z}_2 \;=\; v_2 \;, \tag{5.21}$$

via $z=x$, $v_1=u_1$, $v_2=u_2^3+x_1$. *However, its linear approximation at the origin is* $\dot{x}_1=u_1$, $\dot{x}_2=x_1$, *which is controllable yet not conjugate to* (5.21) *(cf Theorem 4.3).*

[3] A C^0 fibration with fiber $\mathcal{F}$ over $\mathcal{B}$ is a continuous map $g:\mathcal{E}\to\mathcal{B}$ for which every $\xi\in\mathcal{B}$ has a neighborhood $\mathcal{O}$ in $\mathcal{B}$ such that $g^{-1}(\mathcal{O})\subset\mathcal{E}$ is homeomorphic to $\mathcal{O}\times\mathcal{F}$, the so-called *trivializing* homeomorphism $\psi:g^{-1}(\mathcal{O})\to\mathcal{O}\times\mathcal{F}$ being such that $\pi\circ\psi=g$ where $\pi:\mathcal{O}\times\mathcal{F}\to\mathcal{O}$ is the natural projection onto the first factor.

Proof of Theorem 5.9 Let us assume local quasi-smooth linearizability and adopt the notations of Definition 5.6. Without loss of generality, we assume that $\Omega = \mathcal{X} \times \mathcal{U}$ where $\mathcal{X}$ and $\mathcal{U}$ are open neighborhoods of $\bar{x}$ and $\bar{u}$ in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. From (5.2), (5.6), (5.7) and the openness of χ_{II}, we see that $D(x,u)$ coincides in Ω with Ran $\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)^{-1}B$ that clearly does not depend on u. This proves point (1) and defines the distribution Δ_0 in a smooth manner. Consider now another control system $\dot{x} = \hat{f}(x,u)$ defined on Ω by

$$\hat{f}(x,u) \quad = \quad \frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)^{-1} \left[\; A\,\chi_{\mathrm{I}}(x) \; + \; Bu \; \right] . \tag{5.22}$$

On the one hand, this system is smoothly linearizable *via* the diffeomorphism $(x,u) \mapsto (\chi_{\mathrm{I}}(x), u)$ hence it has to meet the necessary and sufficient conditions of Theorem 5.3. On the other hand, thanks to Proposition 5.4, it defines the same distribution Δ_0 in Theorem 5.3 as system (3.1) does in Theorem 5.9, namely Ran $\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)^{-1}B$ for $x \in \mathcal{X}$, and then, since $\hat{f}(x,u) - f(x,u) \in \Delta_0(x)$ for $(x,u) \in \Omega$ by (5.2) and (5.22), an easy induction shows that the higher distributions Δ_k associated to the two systems still coincide. Therefore point (3) follows from Theorem 5.3 applied to $\dot{x} = \hat{f}(x,u)$. To prove point ((2$'$)), let

$$\mathcal{M} = \{ \;\; (x,y) \in \mathcal{X} \times \mathbb{R}^n; \;\; \frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\,y \; - \; A\chi_{\mathrm{I}}(x) \quad \in \quad \operatorname{Ran} B \;\; \} \, .$$

This is a smooth embedded sub-manifold of $\mathcal{X} \times \mathbb{R}^n$ of dimension $n + \operatorname{Rank} B$, this last integer being equal to $n + \widetilde{\nu}(\bar{x},\bar{u})$ by (2) of Proposition 5.8. If we define F as in (5.20), it is clear from (5.2) that

$$F(\mathcal{X} \times \mathcal{U}) \quad \subset \quad \mathcal{M} \, .$$

Now, set $\widetilde{\nu} \; = \; \widetilde{\nu}(\bar{x},\bar{u})$ for simplicity and take some $(m - \widetilde{\nu}) \times m$ matrix C whose rows complement $\widetilde{\nu}$ independent rows of B into a basis of $\mathbb{R}^m$. Pick matrices E_1 and E_2 of appropriate sizes such that

$$E_1\,B \; + \; E_2\,C \quad = \quad I_m \, .$$

By (5.2) we get

$$E_1 \left[\; \frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\,f(x,u) \; - \; A\,\chi_{\mathrm{I}}(x) \; \right] \; + \; E_2 C \chi_{\mathrm{II}}(x,u) \quad = \quad \chi_{\mathrm{II}}(x,u) \, . \tag{5.23}$$

Define

$$\psi : \mathcal{X} \times \mathcal{U} \to \mathcal{M} \times \mathbb{R}^{m-\widetilde{\nu}}$$

by the formula:

$$\psi(x,u) = (x, f(x,u), C\chi_{\mathrm{II}}(x,u)).$$

From (5.23), this mapping has an inverse given by

$$\begin{array}{rcl} \psi^{-1} : \;\; \psi(\mathcal{X} \times \mathcal{U}) & \to & \mathcal{X} \times \mathcal{U} \\ (x,y,z) & \mapsto & \chi^{-1} \left(\; \chi_{\mathrm{I}}(x) \; , \; E_1 \left[\dfrac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\,y \; - \; A\,\chi_{\mathrm{I}}(x) \right] \; + \; E_2\,z \; \right) \end{array}$$

so that ψ defines a homeomorphism from $\mathcal{X}\times\mathcal{U}$ onto its image which is open in $\mathcal{M}\times\mathbb{R}^{m-\widetilde{\nu}}$ by invariance of the domain. Let $\mathcal{O}$ be a neighborhood of $(\bar{x}, f(\bar{x},\bar{u}))$ in $\mathcal{M}$ and $\mathcal{S}$ an open ball centered at $C\chi_{\mathrm{II}}(\bar{x},\bar{u})$ in $\mathbb{R}^{m-\widetilde{\nu}}$ such that $\mathcal{O}\times\mathcal{S}\subset\psi(\mathcal{X}\times\mathcal{U})$, and take $\mathcal{W}=\psi^{-1}(\mathcal{O}\times\mathcal{S})$. Then $F:\mathcal{W}\to F(\mathcal{W})=\mathcal{O}$ is a C^0 fibration with fiber $\mathcal{S}$ and trivializing homeomorphism $\psi:\mathcal{W}\to\mathcal{O}\times\mathcal{S}$. Since $\mathcal{S}$ is homeomorphic to $\mathbb{R}^{m-\widetilde{\nu}}$, condition $(2')$ follows.

Let us now prove sufficiency, assuming without loss of generality that $\mathcal{U}$ is convex. By point (3) we get in particular that Δ_0 is a smooth distribution of constant dimension. Since $\Delta_0(x)=D(x,u)$ for any $u\in\mathcal{U}$ by point (1), we may evaluate this dimension at some (x_M,u_M) where the rank of $\partial f/\partial u$ is locally maximum. Since the rank is locally constant around such points, we have that $D(x_M,u_M)=\operatorname{Ran}\partial f/\partial u(x_M,u_M)$ by Proposition 5.4. Applying the argument to a nested sequence of neighborhoods shrinking to $(\bar{x},\bar{u})$, we conclude in view of (5.17) that the dimension of Δ_0 is necessarily $\widetilde{\nu}$.

Let now $g_1(x),\ldots,g_{\widetilde{\nu}}(x)$ be smooth vector fields forming a basis of $\Delta_0(x)$ at each $x\in\mathcal{X}$; this is possible if we shrink $\mathcal{X}$ by the assumed smoothness of Δ_0. Since $\Delta_0(x)=D(x,u)\supset\operatorname{Ran}\partial f/\partial u\,(x,u)$ for all $(x,u)\in\mathcal{X}\times\mathcal{U}$ by point (1) and Proposition 5.4, there are smooth function $h_{i,j}(x)$ for $1\leq i\leq m$ and $1\leq j\leq\widetilde{\nu}$ such that

$$\frac{\partial f}{\partial u_i}(x,u) \;=\; \sum_{j=1}^{\widetilde{\nu}} h_{i,j}(x,u)g_j(x)\;, \qquad (x,u)\in\mathcal{X}\times\mathcal{U}.$$

Because $\mathcal{U}$ is convex, we may substitute the above relation in:

$$f(x,u)=f(x,\bar{u})+\int_0^1\frac{\partial f}{\partial u}(x,\bar{u}+t(u-\bar{u})).(u-\bar{u})\,dt$$

to obtain smooth functions $\lambda_1(x,u),\ldots,\lambda_{\widetilde{\nu}}(x,u)$ such that

$$f(x,u) \;=\; f(x,\bar{u}) \;+\; \sum_{j=1}^{\widetilde{\nu}}\lambda_j(x,u)g_j(x)\;.$$

Now, if ψ is a trivializing homeomorphism for the C^0 fibration $\mathcal{W}\to F(\mathcal{W})$ provided by point ($(2')$), and if we write $\psi=\psi_1\times\psi_2$ to separate the first $n+\widetilde{\nu}$ components from the $m-\widetilde{\nu}$ remaining ones, we see that

$$(x,u)\mapsto(x,v)=\big(x,\lambda_1(x,u),\ldots,\lambda_{\widetilde{\nu}}(x,u),\psi_2(x,u)\big)$$

yields a continuous injection from $\mathcal{W}$ into $\mathbb{R}^{n+m}$, thus a homeomorphism by invariance of the domain, that quasi-smoothly conjugates system (3.1) to

$$\dot{x} \;=\; f(x,\bar{u}) \;+\; \sum_{j=1}^{\widetilde{\nu}} v_j g_j(x)\;.$$

Granted point (3), one can readily check that this last system satisfies the conditions of Theorem 5.3 for smooth linearizability at $(\bar{x},0,\psi_2(\bar{x},\bar{u}))$. $\Box$

5.3 Non-genericity of (quasi-) smooth linearizability

Smooth linearizability. Except for $m \geq n$ (at least as many controls as states), or for $(n, m) = (2, 1)$, the conditions of Theorem 5.3 impose that a certain number of *equalities* (involving f and its partial derivatives) hold *everywhere* (for example, integrability of a distribution imposes that all Lie brackets be linearly dependent of the original vector fields: a certain number of determinants must be identically zero). This makes these conditions non-generic in any reasonable sense : written in the proper jet spaces, these conditions define a set of infinite co-dimension; also, perturbations of a system that does not satisfy this condition will not satisfy them either, while most perturbations of a system which satisfies them will fail to do so. This is well known; for example, it is shown in [25] that the equivalence classes of any affine control system has infinite codimension in some suitable jet space.

Quasi-smooth linearizability is strictly more general than smooth linearizability, but the gap is quite thin, as shown, from instance by Corollary 5.11. In fact, it imposes that the same equalities hold on an open dense set,although now singularities are allowed. This is no more "generic" than smooth linearizability.

5.4 An open question

Let us examine Conjecture 5.7 in cases that we could not solve, i.e. for a locally quasi-smoothly linearizable system at a point $(\bar{x}, \bar{u})$ such that — see definitions in (5.17) — ν is constant in no neighborhood of $(\bar{x}, \bar{u})$, and $\widetilde{\nu}(\bar{x}, \bar{u}) < m$ (recall that quasi-smooth linearization implies $\widetilde{\nu}$ is locally constant).

In such a case χ_{II} need no longer be smooth, but certain linear combinations of its components must be smooth, namely the entries of $B\chi_{\mathrm{II}}$ as asserted in Proposition 5.8 point 1). Now, letting $\dot{z} = Az + Bv$ be the linear system to which our initial system is quasi-smoothly conjugate, we may assume up to a linear change of variables that the pair (A, B) is in control canonical form (4.7), in which case $B\chi_{\mathrm{II}}$ reduces to the first s_1 components of χ_{II}, s_1 being the rank of B. Thus these first s_1 components are smooth, and we further observe that the last $m - s_1$ components of χ_{II} never appear in equation (5.2) hence can be redefined at will without affecting the conjugation of (3.1) to $\dot{z} = Az + Bv$, provided of course that the newly defined χ is still a homeomorphism. This raises the following question in differential topology which is of interest in its own right and seems to have no answer so far.

Open Question 5.13 *Let O be a neighborhood of the origin in $\mathbb{R}^{p+q}$ and $F : O \to \mathbb{R}^p$ a smooth map. Suppose $G : O \to \mathbb{R}^q$ is a* continuous *map such that $F \times G : O \to \mathbb{R}^p \times \mathbb{R}^q$ is a local homeomorphism at* 0.

Does there exist another neighborhood $O' \subset O$ of the origin and a smooth *map $H : O' \to \mathbb{R}^q$ such that $F \times H : O' \to \mathbb{R}^p \times \mathbb{R}^q$ is also a local homeomorphism at 0 ?*

If the answer to the open question was yes, then Definition 5.6 of quasi-smooth linearizability might equivalently require χ to be smooth because one could set $F = \pi_{n+s_1} \circ \chi$ and smoothly redefine the last $m - s_1$ components of χ.

If the answer to the open question was no, then Definition 5.6 would really be more general than the one obtained by restricting χ to be smooth. Indeed, if F provides a counterexample to the open question, we may consider on $\mathbb{R}^p \times O$ the control system

$$\dot{x} \;=\; F(u) \;, \qquad x \in \mathbb{R}^p,\; u \in \mathbb{R}^{p+q} \tag{5.24}$$

which is locally quasi-smoothly linearizable at the origin since the local homeomorphism

$$(x,u) \;\mapsto\; (z,v) \;=\; (x, F(u), G(u))$$

conjugates (5.24) to

$$\dot{z} \;=\; B\,v, \qquad \text{with } B = (\,I_p\,|0). \tag{5.25}$$

However, no *smooth* homeomorphism

$$\chi\,:\,(x,u) \;\mapsto\; (z,v) \;=\; (\chi_{\mathrm{I}}(x), \chi_{\mathrm{II}}(x,u))$$

exists that quasi-smoothly linearizes (5.24) at 0: if this was the case, by Corollary 4.4 we may assume up to a linear change of variables that χ conjugates (5.24) to (5.25). Then conjugacy would imply

$$\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)F(u) \;=\; B\,\chi_{\mathrm{II}}(x,u)$$

whence in particular

$$F(u) \;=\; \left(\frac{\partial \chi_{\mathrm{I}}}{\partial x}(0)\right)^{-1} B\,\chi_{\mathrm{II}}(0,u),$$

and the last q components of $\chi_{\mathrm{II}}(0,u)$ would yield a smooth H such that $F \times H$ is a local homeomorphism at 0 in R^{p+q}, contrary to the assumption.

6 Topological linearization for control systems

As pointed out in section 5.3, generic systems are *not* smoothly locallly linearizable. If we consider this fact in contrast with Theorem 2.1 on differential equations (without control), it is natural to ask whether removing the differentiability requirement on the conjugacy will allow many more control systems to be locally (topologically) conjugate to a linear controllable one. The theorem to come yields an essentially negative answer to this question. To state the result conveniently, let us first give a formal definition of topological linearizability.

Definition 6.1 *The system* (3.1) *is said to be* locally topologically linearizable at $(\bar{x},\bar{u}) \in \mathbb{R}^{n+m}$ *if it is locally topologically conjugate, in the sense of Definition 3.5, to a linear controllable system.*

Theorem 6.2 *When f is of class C^∞, then system* (3.1) *is locally topologically linearizable at $(\bar{x},\bar{u}) \in \mathbb{R}^{n+m}$ if, and only if, it is locally quasi-smoothly linearizable at $(\bar{x},\bar{u})$.*

Let us make a few comments before proceeding with the proof. The notion of quasi-smooth linearizability was introduced in the previous section for the sole purpose of stating Theorem 6.2 ; the reader is refered to section 5.2 for an extensive discussion of this property, and in particular the gap between quasi-smooth and smooth linearizability.

The above theorem does not say that local topological linearizability is equivalent to local smooth linearizability, and this is actually false as exemplified by system (5.16). However the gap is small, as it appears from Corollaries 5.10 and 5.11; in particular, the former allows one to derive the following corollary to Theorem 6.2 :

Corollary 6.3 *When f is of class C^∞, system* (3.1) *is locally smoothly linearizable at $(\bar{x},\bar{u}) \in \mathbb{R}^{n+m}$ if, and only if, it is locally topologically linearizable at $(\bar{x},\bar{u})$ and the rank of $\partial f/\partial u$ is constant around $(\bar{x},\bar{u})$.*

Proof of Theorem 6.2. Without loss of generality, we suppose that $(\bar{x},\bar{u}) = (0,0)$. Assume there exists a homeomorphism φ from a neighborhood of the origin in $\mathbb{R}^{n+m}$ to an open subset of $\mathbb{R}^{n+m}$ that conjugates system (3.1) to the linear controllable system

$$\dot{z} \;=\; Az \;+\; Bv \tag{6.1}$$

with $z \in \mathbb{R}^n$ and $v \in \mathbb{R}^m$.

First, composing φ with a linear invertible map allows us to suppose that the pair (A,B) is in canonical form (4.7)-(4.8), i.e. that (6.1) can be read

$$\dot{z}_{\sigma_i+k} \;=\; z_{\sigma_{i-1}+k}\;, \quad 2 \le i \le \rho,\; 1 \le k \le s_{i-1}, \tag{6.2}$$

where the integers si and σ_i were defined in (4.5) and (4.6) and where, for notational compactness, we have set :

$$z_{n+k} \;\triangleq\; v_k\;; \tag{6.3}$$

recall here that $s_0 = m$, and notice that $s_1 < m$ may well occur as it simply means that $\operatorname{Rank} B < m$, in which case some of the controls do not appear in the canonical form. With the aggregate notation :

$$Z_j \;\triangleq\; \begin{pmatrix} z_{\sigma_{j+1}+1} \\ \vdots \\ z_{\sigma_j} \end{pmatrix},\; 1 \le j \le \rho-1\;, \qquad Z_0 \;\triangleq\; \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}\;, \tag{6.4}$$

and the matrices J_r^s defined in (4.8), system (6.2) can be rewritten as

$$\begin{array}{rcl} \dot{Z}_{\rho-1} &=& J_{s_{\rho-2}}^{s_{\rho-1}} Z_{\rho-2} \\ \dot{Z}_{\rho-2} &=& J_{s_{\rho-3}}^{s_{\rho-2}} Z_{\rho-3} \\ &\vdots& \\ \dot{Z}_2 &=& J_{s_1}^{s_2} Z_1 \\ \dot{Z}_1 &=& J_{s_0}^{s_1} Z_0 \end{array} \tag{6.5}$$

and is viewed as a control system with state $(Z_{\rho-1}, \ldots, Z_1)$ and control Z_0. We also make the convention, similar to (6.3), that

$$x_{n+k} \triangleq u_k \ , \tag{6.6}$$

and we use for the controls the aggregate notation :

$$X_0 \triangleq \begin{pmatrix} x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} . \tag{6.7}$$

Applying recursively Lemmas F.1 and F.2 (the details of this recursion will be explained shortly, after we have shown how the theorem can be deduced from it), we obtain a *smooth* change of variables around 0 in $\mathbb{R}^n$:

$$(x_1, \ldots, x_n) \ \mapsto \ (X_{\rho-1}, \ldots, X_2, X_1)$$

with $X_i \in \mathbb{R}^{s_i}$ such that, in the new coordinates, system (3.1) reads

$$\begin{array}{rcl} \dot{X}_{\rho-1} & = & F_{\rho-1}(X_{\rho-1}, X_{\rho-2}) \\ \dot{X}_{\rho-2} & = & F_{\rho-2}(X_{\rho-1}, X_{\rho-2}, X_{\rho-3}) \\ & \vdots & \\ \dot{X}_2 & = & F_2(X_{\rho-1}, \ldots, X_1) \\ \dot{X}_1 & = & F_1(X_{\rho-1}, \ldots, X_1, X_0), \end{array} \tag{6.8}$$

and also such that the local homeomorphism Φ that topologically conjugates system (6.8) to system (6.5) at $(0,0)$ is, together with its inverse Ψ, of the triangular form :

$$\begin{array}{rclcrcl} Z_{\rho-1} & = & \Phi_{\rho-1}(X_{\rho-1}) & \quad & X_{\rho-1} & = & \Psi_{\rho-1}(Z_{\rho-1}) \\ Z_{\rho-2} & = & \Phi_{\rho-2}(X_{\rho-1}, X_{\rho-2}) & & X_{\rho-2} & = & \Psi_{\rho-2}(Z_{\rho-1}, Z_{\rho-2}) \\ & \vdots & & & & \vdots & \\ Z_1 & = & \Phi_1(X_{\rho-1}, \ldots, X_1) & & X_1 & = & \Psi_1(Z_{\rho-1}, \ldots, Z_1) \\ Z_0 & = & \Phi_0(X_{\rho-1}, \ldots, X_1, X_0) & & X_0 & = & \Psi_0(Z_{\rho-1}, \ldots, Z_1, Z_0), \end{array} \tag{6.9}$$

where the following three properties hold :

1. Each Φ_k and Ψ_k for $k \geq 1$ is continuously differentiable with respect to X_k and Z_k respectively; in particular, $\partial\Phi_k/\partial X_k$ is invertible throughout the considered neighborhood.
2. For $k \geq 2$, $\text{Rank}\dfrac{\partial F_k}{\partial X_{k-1}}(0, \ldots, 0) \ = \ s_k$, i.e. this rank is maximum, equal to the number of rows.

3. F_1 satisfies

$$
\begin{aligned}
F_1(X_{\rho-1},\ldots,X_1,X_0) \;=\;& F_1(X_{\rho-1},\ldots,X_1,0) \\
&+ \left(\frac{\partial\Phi_1}{\partial X_1}(X_{\rho-1},\ldots,X_1)\right)^{-1} J_m^{s_1} \Bigg(\\
& \Phi_0(X_{\rho-1},\ldots,X_1,X_0) - \Phi_0(X_{\rho-1},\ldots,X_1,0)\Bigg) \,.
\end{aligned}
\tag{6.10}
$$

From the maximum rank assumption on $\partial F_{\rho-1}/\partial X_{\rho-2}$, it is possible to define $Y_{\rho-2}$ whose first $s_{\rho-1}$ entries are those of $F_{\rho-1}(X_{\rho-1},X_{\rho-2})$ and whose remaining $s_{\rho-2}-s_{\rho-1}$ entries are suitable components of $X_{\rho-2}$, in such a way that

$$(X_{\rho-1},\ldots,X_1) \;\mapsto\; (X_{\rho-1},Y_{\rho-2},X_{\rho-3}\ldots,X_1)$$

is a local *smooth* change of coordinates around 0 in $\mathbb{R}^n$. After performing this change of coordinates and setting $Y_{\rho-1}=X_{\rho-1}$ for notational homogeneity, system (6.8) reads

$$
\begin{aligned}
\dot Y_{\rho-1} &= J_{s_{\rho-2}}^{s_{\rho-1}}\, Y_{\rho-2} \\
\dot Y_{\rho-2} &= \widetilde F_{\rho-2}(Y_{\rho-1},Y_{\rho-2},X_{\rho-3}) \\
&\;\vdots \\
\dot X_2 &= \widetilde F_2(Y_{\rho-1},Y_{\rho-2},X_{\rho-3},\ldots,X_1) \\
\dot X_1 &= \widetilde F_1(Y_{\rho-1},Y_{\rho-2},X_{\rho-3},\ldots,X_1,X_0)
\end{aligned}
$$

where the $\widetilde F$'s enjoy the same properties than the F's, in particular the maximality of $\operatorname{Rank}\partial\widetilde F_k/\partial X_{k-1}(0,\ldots,0)$ for $\rho-2\geq k\geq 2$. One may iterate this procedure, limited only by the fact that the maximum rank property mentioned above only holds for $k\geq 2$ but not necessarily for $k=1$. Altogether, this yields a *smooth* local change of coordinates around 0 in $\mathbb{R}^n$:

$$(X_{\rho-1},\ldots,X_1) \;\mapsto\; (Y_{\rho-1},\ldots,Y_1),$$

after which system (6.8) is of the form

$$
\begin{aligned}
\dot Y_{\rho-1} &= J_{s_{\rho-2}}^{s_{\rho-1}}\, Y_{\rho-2} \\
&\;\vdots \\
\dot Y_2 &= J_{s_1}^{s_2}\, Y_1 \\
\dot Y_1 &= F_1(Y_{\rho-1},,\ldots,Y_1,X_0)\;,
\end{aligned}
\tag{6.11}
$$

where we abuse the notation F_1 for simplicity because, although it needs not be the same as in (6.8), this new F_1 enjoys the same property (6.10) for some suitably redefined Φ_1 and Φ_0. Now, we may rewrite (6.10) as

$$F_1(Y_{\rho-1},,\ldots,Y_1,X_0) \;=\; J_m^{s_1}\, H(Y_{\rho-1},,\ldots,Y_1,X_0) \tag{6.11b}$$

where H, in the aggregate notation $Y=(Y_{\rho-1},,\ldots,Y_1)$, is defined by

$$H(Y,X_0) \;=\; \begin{pmatrix}F_1(Y,0)\\0\end{pmatrix} + \begin{pmatrix}\frac{\partial\Phi_1}{\partial Y_1}(Y)^{-1} & 0\\ 0 & I_{m-s_1}\end{pmatrix}\big(\Phi_0(Y,X_0)-\Phi_0(Y,0)\big).$$

Now, because Φ has the triangular structure displayed in (6.9), the map $X_0\mapsto\Phi_0(Y,X_0)$ is injective for fixed $Y=(Y_{\rho-1},\ldots,Y_1)$ in the neighborhood of 0 where it is defined in $\mathbb{R}^m$. Consequently, $(Y,X_0)\mapsto(Y,H(Y,X_0))$ is also injective in the neighborhood of 0 where it is defined in $\mathbb{R}^{n+m}$; since it is continuous, it is a local homeomorphism of $\mathbb{R}^{n+m}$ at $(0,0)$ by invariance of the domain, and then (6.11), (6.11b) make it clear that system (6.8) is locally quasi-smoothly linearizable at this point.

Since (6.8) is smoothly conjugate to the original system (3.1), this proves local quasi-smooth linearizability of the latter hence the theorem.

Let us now detail the argument leading to (6.8)-(6.10). We shall prove, by induction on the non-negative integer ℓ, that the property $\mathcal{P}_\ell$ below is true for $0\le\ell\le\rho-1$, and this will achieve the proof since $\mathcal{P}_{\rho-1}$ is exactly (6.8)-(6.10) if we specialize (6.13) to $i=1$.

Property $\mathcal{P}_\ell$ *: there exists a smooth local change of coordinates around* 0 *in* $\mathbb{R}^n$*, say*

$$(x_1,\ldots,x_n) \;\;\mapsto\;\; (\widehat{X},X_\ell,\ldots,X_2,X_1),$$

with $\widehat{X}\in\mathbb{R}^{\sigma_{\ell+1}}$ *and* $X_i\in\mathbb{R}^{s_i}$ *for* $0\le i\le\ell$ *(if* $\ell=0$ *there are no* X_i*'s beyond* X_0 *whereas if* $\ell=\rho-1$ *there is no* $\widehat{X}$*), after which system* (3.1) *reads:*

$$\begin{array}{rcl}\dot{\widehat{X}} &=& \widehat{F}(\widehat{X},X_\ell)\\ \dot{X}_\ell &=& F_\ell(\widehat{X},X_\ell,X_{\ell-1})\\ &\vdots& \\ \dot{X}_2 &=& F_2(\widehat{X},X_\ell,\ldots,X_1)\\ \dot{X}_1 &=& F_1(\widehat{X},X_\ell,\ldots,X_1,X_0)\end{array}\quad, \tag{6.12}$$

and such that (6.12)*, viewed as a control system with state* $(\widehat{X},X_\ell,\ldots,X_1)$ *and control* X_0*, is locally topologically conjugate at* $(0,0)$ *to system* (6.5) *via a local homeomorphism*

$$(\widehat{X},X_\ell,\ldots,X_1,X_0)\mapsto(Z_{\rho-1},\ldots,Z_0)$$

which is, together with its inverse, of the block triangular form :

$$\begin{array}{rclcrcl}(Z_{\rho-1},\ldots,Z_{\ell+1}) &=& \widehat{\Phi}(\widehat{X}) & \qquad & \widehat{X} &=& \widehat{\Psi}(Z_{\rho-1},\ldots,Z_{\ell+1})\\ Z_\ell &=& \Phi_\ell(\widehat{X},X_\ell) && X_\ell &=& \Psi_\ell(Z_{\rho-1},\ldots,Z_\ell)\\ &\vdots& && &\vdots& \\ Z_1 &=& \Phi_1(\widehat{X},X_\ell,\ldots,X_1) && X_1 &=& \Psi_1(Z_{\rho-1},\ldots,Z_1)\\ Z_0 &=& \Phi_0(\widehat{X},X_\ell,\ldots,X_1,X_0) && X_0 &=& \Psi_0(Z_{\rho-1},\ldots,Z_1,Z_0)\end{array}$$

where Φ_i and Ψ_i are, for $1 \leq i \leq \ell$, continuously differentiable with respect to X_i and Z_i respectively, have an invertible derivative, and satisfy for $1 \leq i \leq \ell$ the relation :

$$\begin{aligned} F_i(\widehat{X}, X_\ell, \ldots, X_i, X_{i-1}) &= F_i(\widehat{X}, X_\ell, \ldots, X_i, 0) \\ + \left(\tfrac{\partial \Phi_i}{\partial X_i}(\widehat{X}, X_\ell, \ldots, X_i)\right)^{-1} J^{s_i}_{s_{i-1}} \Bigg(& \\ \Phi_{i-1}(\widehat{X}, X_\ell, \ldots, X_i, X_{i-1}) &- \Phi_{i-1}(\widehat{X}, X_\ell, \ldots, X_i, 0)\Bigg) \ ; \end{aligned} \tag{6.13}$$

furthermore, the partial homeomorphism

$$(\widehat{X}, X_\ell) \ \mapsto \ (Z_{\rho-1}, \ldots, Z_\ell) \tag{6.14}$$

locally topologically conjugates, at $(0,0) \in I\!R^{\sigma_{\ell+1}+s_\ell}$, the reduced control system

$$\dot{\widehat{X}} \ = \ \widehat{F}(\widehat{X}, X_\ell), \tag{6.15}$$

with state $\widehat{X}$ and input X_ℓ, to the reduced linear control system

$$\begin{aligned} \dot{Z}_{\rho-1} &= J^{s_{\rho-1}}_{s_{\rho-2}} Z_{\rho-2} \ , \\ &\vdots \\ \dot{Z}_{\ell+1} &= J^{s_{\ell+1}}_{s_\ell} Z_\ell \end{aligned} \tag{6.16}$$

wit state $(Z_{\rho-1}, \ldots, Z_{\ell+1})$ and input Z_ℓ.

Granted Lemmas F.1 and F.2 in the appendix, the proof by induction of property $\mathcal{P}_\ell$ is much shorter than its statement. Indeed, $\mathcal{P}_0$ is merely the original assumption on local topological conjugacy of systems (3.1) and (6.5), where the triangular structure (3.10) of the conjugating homeomorphism was taken into account; note in this case that (6.13) is empty and that the reduced system (6.15) is just the original system. Next, supposing that $\mathcal{P}_\ell$ holds for some $\ell \geq 0$, we apply Lemmas F.1 and F.2 to the reduced systems (6.15), (6.16), and to the partial homeomorphism (6.14), with

$$d = \sigma_{\ell+1}, \ r = s_\ell, \ s = s_{\ell+1}, \ U = X_\ell, \ (x_1, \ldots, x_d) = \widehat{X},$$

$$Z^1 = (Z_{\rho-1}, \ldots, Z_{\ell+2}), \ Z^2 = Z_{\ell+1}, \ \text{and } V = Z_\ell,$$

and then, upon renaming $\widetilde{X}^2$ as $X_{\ell+1}$, $\widetilde{f}^2$ as $F_{\ell+1}$, and choosing $\widetilde{X}^1$ to be the new $\widehat{X}$, we get $\mathcal{P}_{\ell+1}$. $\square$

7 Weak Grobman-Hartman theorems for control systems

According to Theorem 6.2 and Section 5.3, generic systems are *not* topologically linearizable; this prevents a general form of the Grobman-Hartman theorem from holding for control systems, at least with definition 3.5 of conjugacy. In retrospect, this may not be too surprising because the Grobman-Hartman theorem for differential equations is about conjugating *flows*, whereas, since the control is an arbitrary function of time whose future values are not determined by past ones, control systems do not have flows, at least of finite dimension. In fact, it is precisely this unpredictability of the future control values that forces the triangular structure of conjugating maps asserted in Proposition 3.7 and ultimately results in the necessary quasi-smoothness of linearizing homeomorphisms claimed by Theorem 6.2.

This section is devoted to positive results on local linearization of control systems. Setting up a stage where a flow can be defined for control systems, either by restricting the inputs or by enlarging the state space to an infinite dimensional one, we derive some analogs to the Grobman-Hartman theorem in that context. These are of course much weaker than topological linearizability in the sense of Definition 6.1, hence do not contradict the previous results.

We consider system (3.1), and suppose that $f(0,0)=0$, *i.e.* we work around an equilibrium point that we choose to be the origin without loss of generality. We assume as usual that f is continuous, and throughout we also make the hypothesis that $\partial f/\partial x(x,u)$ exists and is jointly continuous with respect to (x,u). Subsequently, we single out the linear part of f by consistently setting $A=\frac{\partial f}{\partial x}(0,0)$, so that (3.1) can be rewritten as

$$\begin{array}{rcl} \dot{x} & = & Ax \;+\; P(x,u) \\ \text{with} \quad P(0,0) & = & \dfrac{\partial P}{\partial x}(0,0) \;=\; 0\;. \end{array} \tag{7.1}$$

If in addition f happens to be continuously differentiable with respect to u as well, we set $B=\frac{\partial f}{\partial u}(0,0)$ and we further expand (7.1) into

$$\begin{array}{rcl} \dot{x} & = & Ax \;+\; Bu \;+\; F(x,u) \\ \text{with} \quad F(0,0) & = & \dfrac{\partial F}{\partial x}(0,0) \;=\; \dfrac{\partial F}{\partial u}(0,0) \;=\; 0\;. \end{array} \tag{7.2}$$

Since (7.2) is derived under the stronger hypothesis that f is of class C^1 with respect to both x and u, one would expect stronger results to hold in this case. We want to stress that, deceptively enough, local linearization of (7.2) will turn out to be a consequence of local linearization of (7.1) although the latter was derived without differentiability requirement with respect to u. This is due to the – even more surprising – fact that (7.1) will be locally conjugate to the *non controlled* system $\dot{x}=Ax$, that is to say the influence of the control can be entirely assigned to the linearizing homeomorphism. Compare Theorems 7.4 and 7.6, and see also Remark 7.11.

7.1 An abstract Grobman-Hartman Theorem

Here, we prove an abstract result on the linearization of dynamical systems which implies the results on control systems stated in sections 7.2 and 7.3. The proof closely follows that of the classical Grobman-Hartman theorem for ODEs as given by Hartman in [8, chap. IX, sect. 4, 7, 8, 9], and we tried to stick to his notations as much as possible. Nevertheless, we provide a detailed argument because the modifications needed to handle the dynamics of the control are not completely straightforward. Like [8], we state Theorem 7.1 as a *global* linearizability property for a linear equation perturbed by a suitably normalized additive term. In sections 7.2 and 7.3, we shall use this result to derive local linearizability results for systems that locally coincide with a normalized one.

Let us mention in passing that the Grobman-Hartman Theorem for "random dynamical systems" given in [6] is similar in spirit to Theorem 7.1 : there, the set $\mathcal{E}$ of control parameters is a probability space instead of a topological space, and the conjugating transformation H is only required to be measurable with respect to $\zeta \in \mathcal{E}$ but need not be continuous. Both can be viewed as Grobman-Hartman Theorem "with parameters".

The setting is as follows. We consider a topological space $\mathcal{E}$ endowed with a one-parameter group of homeomorphisms $(\mathcal{S}_\tau)_{\tau\in\mathbb{R}}$. The space $\mathcal{E}$ is to be regarded as an abstract collection of input-producing events for a control system, these events being themselves subject to the dynamics of the flow $\mathcal{S}_\tau$. To describe the action of such an event on system (7.1), we simply let ζ enter as a parameter in the differential equation describing the evolution of the state variable x where, as is customary in this section, the linear term at the origin has been singled out:

$$\dot{x} \;=\; Ax \;+\; G(x,\zeta,t)\;, \tag{7.3}$$

and where $G : \mathbb{R}^n \times \mathcal{E} \times \mathbb{R} \to \mathbb{R}^n$ is assumed to be measurable with respect to t for fixed x, ζ, and of class C^1 with respect to x for fixed ζ, t. To ensure the compatibility between the dynamics of ζ and that of x (see (7.6) below), we also require the condition

$$G(x,\mathcal{S}_\tau(\zeta),t) = G(x,\zeta,t+\tau) \tag{7.4}$$

to hold for all $(x,\zeta,\tau,t) \in \mathbb{R}^n \times \mathcal{E} \times \mathbb{R} \times \mathbb{R}$. Now, if we suppose that to each $(x,\zeta) \in \mathbb{R}^n \times \mathcal{E}$ there is a locally integrable function $\phi_{x,\zeta} : \mathbb{R} \to \mathbb{R}^+$ satisfying $G(x,\zeta,t) \leq \phi_{x,\zeta}(t)$ for all $t \in \mathbb{R}$, and that to each $\zeta \in \mathcal{E}$ there is a locally integrable function $\psi_\zeta : \mathbb{R} \to \mathbb{R}^+$ satisfying $\partial G/\partial x\,(x,\zeta,t) \leq \psi_\zeta(t)$ for all $(x,t) \in \mathbb{R}^n \times \mathbb{R}$, then for each $\zeta \in \mathcal{E}$ the solution to (7.3) with initial condition $x(0) = x_0 \in \mathbb{R}^n$ uniquely exists for all $t \in \mathbb{R}$, *cf.* [22, Theorem 54, Proposition C.3.4, Proposition C.3.8]. Subsequently, denoting by

$$\widehat{x}(\tau,x_0,\zeta) \tag{7.5}$$

the value of this solution at time $t = \tau$, it follows from (7.4) that

$$\widehat{x}(t+\tau,x_0,\zeta) = \widehat{x}(t,\widehat{x}(\tau,x_0,\zeta),\mathcal{S}_\tau(\zeta)) \tag{7.6}$$

and thus

$$\widehat{\Phi}_t(x_0,\zeta) \;=\; (\,\widehat{x}(t,x_0,\zeta)\,,\,\mathcal{S}_t(\zeta)\,) \tag{7.7}$$

defines a flow on $\mathbb{R}^n \times \mathcal{E}$, the group property being a consequence of (7.6) and of the group property of $\mathcal{S}_\tau$. We call $(\widehat{\Phi}_t)_{t\in\mathbb{R}}$ the flow of system (7.3).

We also define the partially linear flow L_t by the formula:

$$L_t(x_0, \zeta) \;=\; (\, e^{tA} x_0 \,,\, \mathcal{S}_t(\zeta)\,)\;; \tag{7.8}$$

it is the flow of (7.3) when $G = 0$, and the whole point in this subsection is to give conditions on G for $\widehat{\Phi}_t$ and L_t to be topologically conjugate over $\mathbb{R}^n \times \mathcal{E}$.

We will assume throughout that the $n \times n$ matrix A is hyperbolic, hence it is similar to a block diagonal one:

$$A \sim \begin{pmatrix} A_e & 0 \\ 0 & A_l \end{pmatrix} \tag{7.9}$$

where A_e and A_l are $e\times e$ and $l\times l$ real matrices, with $e+l = n$, whose eigenvalues have strictly negative and strictly positive real parts respectively. Now, there exist Euclidean norms on $\mathbb{R}^e$ and $\mathbb{R}^l$ for which e^{A_e} and e^{-A_l} are strict contractions, because their eigenvalues have modulus strictly less than 1 and any square complex matrix is similar to an upper triangular one having the eigenvalues of the original matrix as diagonal entries while the remaining entries are arbitrarily small, see *e.g.* [1, ch.3, sec.22.4, Lemma 4]. Therefore, combining (7.9) with a suitable linear change of variable on each factor in $\mathbb{R}^n = \mathbb{R}^e \times \mathbb{R}^l$, we can write

$$A = E^{-1} \begin{pmatrix} P & 0 \\ 0 & Q \end{pmatrix} E, \tag{7.10}$$

where E is some nonsingular $n\times n$ real matrix while P and Q are $e\times e$ and $l \times l$ real matrices such that e^P and e^{-Q} are strict contractions for the standard Euclidean norm:

$$c \;\triangleq\; \|e^P\|_{\mathrm{O}} < 1 \quad \text{and} \quad \frac{1}{d} \;\triangleq\; \|e^{-Q}\|_{\mathrm{O}} < 1\;, \tag{7.11}$$

where $\|.\|_{\mathrm{O}}$ designates the familiar operator norm of a matrix. Subsequently, we define the real numbers

$$b_1 \;\triangleq\; \|e^{-P}\|_{\mathrm{O}} + \|e^{-Q}\|_{\mathrm{O}} \;=\; \frac{1}{d} + \|e^{-P}\|_{\mathrm{O}}\,, \tag{7.12}$$

$$c_1 \;\triangleq\; \left\|EAE^{-1}\right\|_{\mathrm{O}} \;=\; \max\{\,\|P\|_{\mathrm{O}}\,,\,\|Q\|_{\mathrm{O}}\}. \tag{7.13}$$

Besides the operator norm, we shall make use of another norm on real matrices, namely the Frobenius norm $\|.\|_{\mathrm{F}}$ which is the square root of the sum of the squares of the entries. Let us record the elementary inequalities, valid for any two real square matrices M, N :

$$\|M\|_{\mathrm{O}} \le \|M\|_{\mathrm{F}}\,,\; \|MN\|_{\mathrm{F}} \le \min\{\|M\|_{\mathrm{O}}\|N\|_{\mathrm{F}}, \|M\|_{\mathrm{F}}\|N\|_{\mathrm{O}}\}. \tag{7.14}$$

As usual, we keep the symbol $\|.\|$ to indicate the standard Euclidean norm on $\mathbb{R}^j$ irrespectively of j. Now, our main result is the following:

Theorem 7.1 *Let the hyperbolic matrix A and the numbers c, d, b_1 and c_1 be as in* (7.10), (7.11), (7.12) *and* (7.13). *Assume that the topological space $\mathcal{E}$, its one-parameter group of homeomorphisms $(\mathcal{S}_\tau)$, and the map $G : \mathbb{R}^n \times \mathcal{E} \times \mathbb{R} \to \mathbb{R}^n$ satisfy the following conditions :*

- *Equation* (7.4) *holds for all* $(x,\zeta,\tau,t) \in \mathbb{R}^n \times \mathcal{E} \times \mathbb{R} \times \mathbb{R}$.
- *For fixed* $\zeta \in \mathcal{E}$, *the map* $\tau \mapsto \mathcal{S}_\tau(\zeta)$ *is Borel measurable* $\mathbb{R} \to \mathcal{E}$, *that is to say the inverse image of an open subset of* $\mathcal{E}$ *is measurable in* $\mathbb{R}$.
- *The map* $x \mapsto G(x,\zeta,t)$ *is continuously differentiable* $\mathbb{R}^n \to \mathbb{R}^n$ *for fixed* $(\zeta,t) \in \mathcal{E}\times\mathbb{R}$, *the map* $t \mapsto G(x,\zeta,t)$ *is measurable* $\mathbb{R} \to \mathbb{R}^n$ *for fixed* $(x,\zeta) \in \mathbb{R}^n \times \mathcal{E}$, *and to each* $\zeta \in \mathcal{E}$ *there are locally integrable functions* $\phi_\zeta, \psi_\zeta : \mathbb{R} \to \mathbb{R}^+$ *such that, for all* $(x,t) \in \mathbb{R}^n \times \mathbb{R}$, *one has :*

$$\|G(x,\zeta,t)\| \;\leq\; \phi_\zeta(t)\ , \qquad \|\frac{\partial G}{\partial x}(x,\zeta,t)\|_{\mathrm{F}} \;\leq\; \psi_\zeta(t)\ . \tag{7.15}$$

- *Defining the flow* $\widehat{x}$ *of* (7.3) *as in* (7.5), *the map* $(x_0,\zeta) \mapsto \widehat{x}(t,x_0,\zeta)$ *is continuous* $\mathbb{R}^n \times \mathcal{E} \to \mathbb{R}^n$ *for fixed* $t \in \mathbb{R}$.
- *There are real numbers* $M > 0$ *and* $\eta > 0$ *such that*

$$\forall \zeta \in \mathcal{E} \quad , \quad \|\phi_\zeta\|_{L^1([0,1])} \;\leq\; M\ , \tag{7.16}$$

$$\|\psi_\zeta\|_{L^1([0,1])} \;\leq\; \eta\ ; \tag{7.17}$$

  Moreover, the number η *in* (7.17) *is so small that, putting*

$$\theta \triangleq \eta\, \|E\|_{\mathrm{O}}\, \|E^{-1}\|_{\mathrm{O}}$$

  and then

$$\alpha_1 \;\triangleq\; \theta e^{c_1}\left(1 + e^{\theta + c_1}\,(\theta + c_1)\right),$$

  one has

$$0 < b_1\alpha_1 < 1 \quad \textit{and} \quad \alpha_1(1 + 1/d) + \max(c, 1/d) < 1\ . \tag{7.18}$$

Then, there exists a homeomorphism

$$\mathcal{H} :\; \mathbb{R}^n \times \mathcal{E} \to \mathbb{R}^n \times \mathcal{E}$$

of the form

$$(x,\zeta) \quad \mapsto \quad \mathcal{H}(x,\zeta) = (H(x,\zeta),\zeta),$$

that conjugates $\widehat{\Phi}_t$ *defined in* (7.7) *to the partially linear flow* (7.8), *namely* $\mathcal{H} \circ \widehat{\Phi}_t = L_t \circ \mathcal{H}$ *or, equivalently,*

$$H(\,\widehat{\Phi}_t(x,\zeta)\,) \;\;=\;\; e^{tA} H(x,\zeta) \tag{7.19}$$

for all $(t,x,\zeta) \in \mathbb{R} \times \mathbb{R}^n \times \mathcal{E}$.

To establish Theorem 7.1, we shall rely on two lemmas. The first one runs parallel to [8, chap. IX, lemma 8.3], and gives us sufficient conditions for perturbations of a map $(x, \zeta) \mapsto (Lx, \mathcal{S}_\tau(\zeta))$ to be topologically conjugate on $\mathbb{R}^n \times \mathcal{E}$, when τ is fixed and the linear map $L : \mathbb{R}^n \to \mathbb{R}^n$ is the product of a dilation and a contraction. This lemma is the mainspring of the proof, in that it will provide us with the desired conjugating $\mathcal{H}$ when applied to the flows (7.7) and (7.8) evaluated at $t = 1$ (this arbitrary value comes from the normalization of the constants c and d through (7.11)). The proof of the lemma is similar to that of [8, chap. IX, lemma 8.3], except that we need to keep track more carefully of uniqueness and continuity issues here; it uses the shrinking lemma on Lipschitz-small perturbations of hyperbolic linear maps, a classical device to build conjugating homeomorphisms that has many other applications, see [8, chap. IX, notes]. The reader will notice that the statement of the lemma redefines the constants c, d, b_1, and α_1 that were already fixed in the statement of Theorem 7.1. We allow ourself this minor incorrection, because we feel it helps following the argument since the lemma will be applied precisely with the previously defined constants.

Lemma 7.2 *Let us be given a homeomorphism $\mathcal{T} : \mathcal{E} \to \mathcal{E}$ and two non-singular real matrices C, D of size $e \times e$ and $l \times l$ respectively, such that $c = \|C\| < 1$ and $\frac{1}{d} = \|D^{-1}\| < 1$.*

For $i = 1, 2$, let $Y_i : \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E} \to \mathbb{R}^e$ and $Z_i : \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E} \to \mathbb{R}^l$ be two pairs of bounded continuous functions satisfying

$$\max\{\|\Delta Y_i\|, \|\Delta Z_i\|\} \leq \alpha_1(\|\Delta y\| + \|\Delta z\|), \tag{7.20}$$

where ΔY_i and ΔZ_i stand respectively for $Y_i(y + \Delta y, z + \Delta z, \zeta) - Y_i(y, z, \zeta)$ and $Z_i(y + \Delta y, z + \Delta z, \zeta) - Z_i(y, z, \zeta)$, and where α_1 is a constant such that, if we put $a = \|C^{-1}\|$ and $b_1 = a + 1/d$, then $0 < b_1\alpha_1 < 1$ and $\alpha_1(1 + 1/d) + \max(c, 1/d) < 1$. If we define for $i = 1, 2$ the maps

$$\begin{array}{cccc} T_i : & \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E} & \to & \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E} \\ & (y, z, \zeta) & \mapsto & (Cy + Y_i(y, z, \zeta), Dz + Z_i(y, z, \zeta), \mathcal{T}(\zeta)), \end{array}$$

then there exists a unique map $R_0 : \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E} \to \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E}$ of the form

$$R_0(y, z, \zeta) = (H_0(y, z, \zeta), \zeta) \tag{7.21}$$

such that:

- *$H_0(y, z, \zeta) - (y, z)$ is bounded on $\mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E}$,*
- *one has the commuting relation:*

$$R_0 T_1 = T_2 R_0. \tag{7.22}$$

Moreover, R_0 is then necessarily a homeomorphism of $\mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E}$.

The second lemma that we need in order to prove Theorem 7.1 is of technical nature and ensures that, under the hypotheses stated in that proposition, we can indeed apply Lemma 7.2 to the flow (7.7) evaluated at $t=1$. Recalling from (7.5) the definition of $\widehat{x}$, it will be convenient to define a map $\Xi : I\!R \times I\!R^n \times \mathcal{E} \to I\!R^n$ by the equation :

$$\widehat{x}(t, x_0, \zeta) \quad = \quad \exp(tA)\, x_0 + \Xi(t, x_0, \zeta)\ . \tag{7.23}$$

Thus the map Ξ capsulizes the deviation of the flow of (7.3) from the flow of the linearized equation $\dot{x} = Ax$.

Lemma 7.3 *Under the assumptions of Theorem 7.1, the map Ξ defined by* (7.23) *is bounded on* $[0,1] \times I\!R^n \times \mathcal{E}$*, it is of class C^1 with respect to x_0 for fixed t, ζ, and it satisfies, for all* (t, x_0, ζ) *in* $[0,1] \times I\!R^n \times \mathcal{E}$*, the inequality :*

$$\|\frac{\partial \Xi}{\partial x_0}(t, x_0, \zeta)\|_{\mathrm{F}} \;\leq\; \eta\, e^{\|A\|_{\mathrm{O}}} \left(1 + e^{\eta + \|A\|_{\mathrm{O}}} \,(\eta + \|A\|_{\mathrm{O}})\right). \tag{7.24}$$

Assuming Lemma 7.2 and Lemma 7.3 for a while, let us proceed immediately with the proof of Theorem 7.1.

Proof of Theorem 7.1. Performing on $I\!R^n$ the change of variables $x \mapsto E\,x$ and taking (7.14) into account, we may assume upon replacing M by $M\|E\|_{\mathrm{O}}$ in (7.16) and η by θ in (7.17) that $E = I_n$, the identity matrix of size n. Then $c_1 = \|A\|_{\mathrm{O}}$ and the right-hand side of (7.24) is just α_1. Moreover (7.10) expresses that A assumes a block-diagonal form, according to which we block-decompose the flow $\widehat{\Phi}_t(x_0, \zeta)$ defined by (7.7)) into

$$\begin{pmatrix} y_0 \\ z_0 \\ \zeta \end{pmatrix} \quad \mapsto \quad \begin{pmatrix} e^{tP} y_0 + Y(t, y_0, z_0, \zeta) \\ e^{tQ} z_0 + Z(t, y_0, z_0, \zeta) \\ \mathcal{S}_t(\zeta) \end{pmatrix} \tag{7.25}$$

where $(y_0^T, z_0^T)^T$ is the natural partition of $x_0 \in I\!R^n \sim I\!R^e \times I\!R^l$, and where Y and Z are respectively the first e and the last l components of the map Ξ defined in (7.23). Still taking into account the block decomposition induced by (7.10) where $E = I_n$, the partially linear flow L_t defined by (7.8) in turn splits into

$$\begin{array}{rccl} L_t : & I\!R^e \times I\!R^d \times \mathcal{E} & \to & I\!R^e \times I\!R^d \times \mathcal{E} \\ & (y_0, z_0, \zeta) & \mapsto & (\exp(Pt)\, y_0, \exp(Qt)\, z_0, \mathcal{S}_t(\zeta)). \end{array}$$

We shall apply Lemma 7.2 with $\mathcal{T} = \mathcal{S}_1$ to $T_1 = \widehat{\Phi}_1$ and $T_2 = L_1$, that is to say we choose $C = e^P$, $D = e^Q$, $Y_2 = 0$, $Z_2 = 0$, and we define Y_1 and Z_1 by $Y_1(y,z,\zeta) = Y(1,y,z,\zeta)$ and $Z_1(y,z,\zeta) = Z(1,y,z,\zeta)$ where Y, Z are as in (7.25). The hypotheses on C and D are satisfied by (7.11), while the hypotheses on Y_2 and Z_2 are trivially met. As to Y_1 and Z_1, we observe that:
- their continuity, i.e. the continuity of $(x_0, \zeta) \mapsto \Xi(1, x_0, \zeta)$, follows *via* (7.23) from the continuity of $(x_0, \zeta) \mapsto \widehat{x}(1, x_0, \zeta)$ which is part of the hypotheses (see point 4 in the statement

of the proposition);
- their boundedness, i.e. the boundedness of $(x_0, \zeta) \mapsto \Xi(1, x_0, \zeta)$, follows from Lemma 7.3;
- the inequalities on the Lipschitz constants of Y_1 and Z_1 required in Lemma 7.2 follow from the mean-value theorem and Lemma 7.3, equation (7.24), granted (7.18), (7.14), and the triangle inequality.

Therefore Lemma 7.2 does apply, providing us with a homeomorphism of $\mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E} = \mathbb{R}^n \times \mathcal{E}$ of the form $R_0 = H_0 \times \mathrm{id}$, which is such that $H_0(x, \zeta) - x$ is bounded on $\mathbb{R}^n \times \mathcal{E}$ and, in addition, such that

$$R_0 \circ \widehat{\Phi}_1 \;=\; L_1 \circ R_0. \tag{7.26}$$

Equation (7.26) expresses that H_0 conjugates the flow $\widehat{\Phi}_t(x, \zeta))$ to the partially linear flow L_t at time $t = 1$, whereas we want these flows to be conjugate at any time t. For this, we use the same averaging trick (originally due to S. Sternberg) as in [8, chap. IX, sec. 9], namely we define $H : \mathbb{R}^n \times \mathcal{E} \to \mathbb{R}^n$ by the integral formula:

$$H(x, \zeta) = \int_0^1 e^{-rA} H_0(\widehat{\Phi}_r(x, \zeta))\, dr \tag{7.27}$$

where H_0, being the first factor of R_0, must satisfy by virtue of (7.26):

$$H_0(\widehat{\Phi}_1(x, \zeta)) \;=\; e^A H_0(x, \zeta) \tag{7.28}$$

We need of course show that (7.27) is well-defined. Firstly, let us check that the integrand is a measurable function of r. As H_0 is continuous $\mathbb{R}^n \times \mathcal{E} \to \mathbb{R}^n$, this reduces to showing that the map

$$r \mapsto \widehat{\Phi}_r(x, \zeta) = (\, \widehat{x}(r, x, \zeta)\, ,\, \mathcal{S}_r(\zeta)\,) \tag{7.29}$$

is measurable $\mathbb{R} \mapsto \mathbb{R}^n \times \mathcal{E}$. Now, the map $r \mapsto \widehat{x}(r, x, \zeta)$ is *a fortiori* measurable since it is absolutely continuous, and the map $r \mapsto \mathcal{S}_r(\zeta)$ is also measurable by assumption (see point 2 in the statement of the proposition). Hence the inverse image under (7.29) of an open rectangle is measurable in $\mathbb{R}$. But any open subset of $\mathbb{R}^n \times \mathcal{E}$ is a countable union of open rectangles because $\mathbb{R}^n$ has a countable basis of open neighborhoods, and this establishes the measurability of (7.29). Secondly, the integrand in (7.27) is bounded, for $\|H_0(\widehat{\Phi}_r(x, \zeta)) - \widehat{x}(r, x, \zeta)\|$ is majorized uniformly with respect to r, x, and ζ since $H_0(x, \zeta) - x$ is bounded on $\mathbb{R}^n \times \mathcal{E}$ by the properties of R_0, while the continuous function $r \mapsto \widehat{x}(r, x, \zeta)$ is bounded for fixed x and ζ on the compact set $[0, 1]$. Therefore, the integral on the right-hand side of (7.27) indeed exists.
Observe now that $H(x, \zeta) - x$ is also bounded on $\mathbb{R}^n \times \mathcal{E}$. Indeed, by definition of $\widehat{\Phi}_r$ *via* (7.7) and of Ξ *via* (7.23), we can write

$$\begin{aligned} H(x, \zeta) - x \;&=\; \int_0^1 e^{-rA} \big(H_0(\, \widehat{x}(r, x, \zeta)\, ,\, \mathcal{S}_r(\zeta)\,) - \widehat{x}(r, x, \zeta) \,\big)\, dr \\ &+\; \int_0^1 e^{-rA}\, \Xi(r, x, \zeta)\, dr, \end{aligned} \tag{7.30}$$

and since both integrals on the right-hand side are bounded (the first because $H_0(x,\zeta)-x$ is bounded on $I\!R^n \times \mathcal{E}$ and the second because Ξ is bounded on $[0,1]\times I\!R^n \times \mathcal{E}$ by Lemma 7.3, we get the desired boundedness of $H(x,\zeta)-x$. Next, *we claim that* (7.19) *holds*, and once we have proved this the proposition will follow because, specializing (7.19) to $t=1$, we shall conclude by the uniqueness part of Lemma 7.2 that $H\times \mathrm{id} = R_0$ and therefore that R_0, which is a homeomorphism of $I\!R^n \times \mathcal{E}$ with the desired form, will meet $R_0 \circ \widehat{\Phi}_t = L_t \circ R_0$, not just for $t=1$ as we knew already but in fact for all t. Thus it will be possible to take $\mathcal{H} = R_0$.

To establish the claim, we use the group property of the flow to write

$$e^{-tA} H(\,\widehat{\Phi}_t(x,\zeta)\,) \;\;=\;\; \int_0^1 e^{-(t+r)A} H_0(\widehat{\Phi}_{t+r}(x,\zeta)\,)\,dr,$$

and we set $t+r=\tau$ to convert the above integral into

$$\int_t^{t+1} e^{-\tau A} H_0(\widehat{\Phi}_\tau(x,\zeta)\,)\,\mathrm{d}\tau \;\;=\;\; \int_t^1 \ldots\;\mathrm{d}\tau + \int_1^{t+1} \ldots\;\mathrm{d}\tau, \tag{7.31}$$

where the dots indicate that the integrand is repeated in each integral. Now, putting $\lambda = \tau - 1$, the last integral in the right-hand side becomes

$$\int_0^t e^{-(\lambda+1)A} H_0(\widehat{\Phi}_{\lambda+1}(x,\zeta)\,)\,d\lambda \;\;=\;\; \int_0^t e^{-\lambda A} H_0(\widehat{\Phi}_\lambda(x,\zeta)\,)\,d\lambda,$$

where we have used the group property of the flow again together with (7.28). Plugging this into (7.31), we recover back $\int_0^1 e^{-tA} H_0(\widehat{\Phi}_t(x,\zeta)\,)dt$ on the right-hand side, so that finally $e^{-tA} H \circ \widehat{\Phi}_t \;=\; H$ *as claimed.* $\square$

Let us now tie the loose ends in the proof of Theorem 7.1 by establishing Lemma 7.3 and Lemma 7.2.

Proof of Lemma 7.3 From (7.3) and (7.23), we see that $t \mapsto \Xi(t,x_0,\zeta)$ is the solution to

$$\dot{\xi}(t) = A\xi(t) + G\left(\xi(t) + e^{tA}x_0\,,\,\zeta\,,\,t\right)$$

with initial condition $\xi(0)=0$. Since $\|G(x,\zeta,t)\|$ is bounded by $\phi_\zeta(t)$ with $\|\phi_\zeta\|_{L^1([0,1])} \le M$ by (7.15) and (7.16), we get by integration that

$$\|\xi(t)\| \le M + \|A\|_{\mathrm{O}} \int_0^t \|\xi(s)\|\,d|s|, \qquad t\in[0,1],$$

so by the Bellman-Gronwall lemma (*cf* Lemma B.1 in Appendix (B)) :

$$\|\xi(t)\| \le M\left(1+|t|\,\|A\|_{\mathrm{O}}\,e^{|t|\|A\|_{\mathrm{O}}}\right), \qquad t\in[0,1].$$

This entails that Ξ is bounded on $[0,1]\times \mathbb{R}^n \times \mathcal{E}$.

To prove (7.24), we consider for fixed x_0, ζ the matrix-valued function $R(t) = \frac{\partial \widehat{x}}{\partial x_0}(t, x_0, \zeta)$, whose existence and continuity with respect to x_0 for fixed t, ζ depend on (7.15), (7.16) and (7.17) (*cf* Proposition B.2 in Appendix B), inducing in turn the existence and continuity with respect to x_0 of $Q(t) = \frac{\partial \Xi}{\partial x_0}(t, x_0, \zeta)$ *via* (7.23). The variational equation for $\frac{\partial \widehat{x}}{\partial x_0}$ (see again Proposition B.2 in Appendix B) yields :

$$\dot{R}(t) \;=\; \left[A + \frac{\partial G}{\partial x}\big(\widehat{x}(t, x_0, \zeta), \zeta, t\big)\right] R(t)\;, \quad R(0) \;=\; I_n\;,$$

and, since $R(t) = Q(t) + e^{tA}$ by (7.23), we have that

$$\dot{Q}(t) = \left[A + \frac{\partial G}{\partial x}\big(\widehat{x}(t, x_0, \zeta), \zeta, t\big)\right] Q(t) + \frac{\partial G}{\partial x}\big(\widehat{x}(t, x_0, \zeta), \zeta, t\big)\, e^{tA}\,,\ Q(0) = 0.$$

Put $\rho(t) = \|Q(t)\|_{\mathrm{F}}$. Due to the definition of the Frobenius norm, $\rho(t)$ is locally absolutely continuous and, by the Cauchy-Schwarz inequality, one has $\dot{\rho}(t) \leq \|\dot{Q}(t)\|_{\mathrm{F}}$. Thus, the differential equation satisfied by $Q(t)$ together with (7.15) yield :

$$\dot{\rho} \;\leq\; (\psi_\zeta(t) + \|A\|_{\mathrm{O}})\, \rho(t) \;+\; \psi_\zeta(t)\, e^{|t|\,\|A\|_{\mathrm{O}}}, \qquad \rho(0) = 0,$$

where we have used (7.14) and the elementary fact that $\|e^{tA}\|_{\mathrm{O}} \leq e^{|t|\,\|A\|_{\mathrm{O}}}$. Integrating this inequality and applying the Bellman-Gronwall lemma (*cf* Lemma B.1 in Appendix (B)) while taking (7.17) into account leads us to

$$\rho(t) \;\;\leq\;\; \eta\, e^{|t|\,\|A\|_{\mathrm{O}}} \left(1 + e^{\eta + |t|\,\|A\|_{\mathrm{O}}}\,(\eta + |t|\,\|A\|_{\mathrm{O}})\right), \qquad t \in [0,1].$$

By definition of ρ, this implies (7.24). $\square$

Proof of Lemma 7.2 If we endow $\mathbb{R}^e \times \mathbb{R}^l$ with the norm $\|(y,z)\| = \|y\| + \|z\|$, it follows from (7.20) that, for fixed $(y, z, \zeta) \in \mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E}$, the map $T_{y,z,\zeta} : \mathbb{R}^e \times \mathbb{R}^l \to \mathbb{R}^e \times \mathbb{R}^l$ defined by

$$T_{y,z,\zeta}(y', z') = (C^{-1}y, D^{-1}z) - \big(C^{-1}Y_1(y', z', \zeta)\,,\, D^{-1}Z_1(y', z', \zeta)\big)$$

is a shrinking map with shrinking constant $b_1\alpha_1 < 1$, whose fixed point is the unique $(\bar{y}, \bar{z}) \in \mathbb{R}^e \times \mathbb{R}^l$ satisfying $T_1(\bar{y}, \bar{z}, \zeta) = (y, z, \mathcal{T}(\zeta))$. In addition, it holds that $(\bar{y}, \bar{z}) = \lim_{k\to\infty} T^k_{y,z,\zeta}(y', z')$ for any (y', z'), and this classically implies that $(\bar{y}, \bar{z})$ is continuous with respect to y, z, and ζ. Indeed, the continuity of Y_1 and Z_1 entails that $T_{y,z,\zeta}(y', z')$ is continuous with respect to y, z and ζ for fixed y', z'. Therefore, if we write $\bar{y}(y, z, \zeta)$, $\bar{z}(y, z, \zeta)$ to emphasize the functional dependence, and if we choose y_0, z_0, ζ_0 together with $\varepsilon > 0$, there is a neighborhood $\mathcal{V}_0$ of (y_0, z_0, ζ_0) in $\mathbb{R}^e \times \mathbb{R}^l \times \mathcal{E}$ such that $(y, z, \zeta) \in \mathcal{V}_0$ implies :

$$\begin{aligned}
&\|T_{y,z,\zeta}(\bar{y}(y_0, z_0, \zeta_0), \bar{z}(y_0, z_0, \zeta_0)) - (\bar{y}(y_0, z_0, \zeta_0), \bar{z}(y_0, z_0, \zeta_0))\| \\
&= \|T_{y,z,\zeta}(\bar{y}(y_0, z_0, \zeta_0), \bar{z}(y_0, z_0, \zeta_0)) - T_{y_0,z_0,\zeta_0}((\bar{y}(y_0, z_0, \zeta_0), \bar{z}(y_0, z_0, \zeta_0))\| < \varepsilon.
\end{aligned}$$

Consequently, for $(y,z,\zeta)\in\mathcal{V}_0$, we have by the shrinking property that

$$\begin{aligned}&\left\|\left(\bar{y}(y,z,\zeta),\bar{z}(y,z,\zeta)\right)-\left(\bar{y}(y_0,z_0,\zeta_0),\bar{z}(y_0,z_0,\zeta_0)\right)\right\|\\ &=\|\lim_{k\to\infty}T^k_{y,z,\zeta}((\bar{y}(y_0,z_0,\zeta_0),\bar{z}(y_0,z_0,\zeta_0))-(\bar{y}(y_0,z_0,\zeta_0),\bar{z}(y_0,z_0,\zeta_0))\|\\ &\le\sum_{k=0}^{\infty}\|T^{k+1}_{y,z,\zeta}((\bar{y}(y_0,z_0,\zeta_0),\bar{z}(y_0,z_0,\zeta_0))-T^k_{y,z,\zeta}((\bar{y}(y_0,z_0,\zeta_0),\bar{z}(y_0,z_0,\zeta_0))\|\\ &\le\frac{\varepsilon}{1-b_1\alpha_1}\end{aligned}$$

which implies the desired continuity. Therefore, $(x,y)\mapsto(\bar{y}(y,z,\zeta),\bar{z}(y,z,\zeta))$ is, for fixed ζ, the inverse of the concatenation of the first two components of T_1, and it is continuous with respect to (x,y), and to ζ. Moreover, we see from the definition of $T_{y,z,\zeta}$ and the fixed point property of $\bar{y},\bar{z}$ that

$$(\bar{y},\bar{z})=(C^{-1}y,D^{-1}z)-(C^{-1}Y_1(\bar{y},\bar{z},\zeta)\,,\,D^{-1}Z_1(\bar{y},\bar{z},\zeta))$$

and, since Y_1 and Z_1 are continuous and bounded, this makes for a relation of the form

$$(\bar{y}(y,z,\zeta),\bar{z}(y,z,\zeta))=(C^{-1}y+\hat{Y}_1(y,z,\zeta),D^{-1}z+\hat{Z}_1(y,z,\zeta))$$

where $\hat{Y}_1$, $\hat{Z}_1$ are in turn continuous and bounded on $I\!R^e\times I\!R^l\times\mathcal{E}$ with values in $I\!R^e$ and $I\!R^l$ respectively. All this yields the existence of an inverse for the map T_1 itself, namely

$$T_1^{-1}(y,z,\zeta)=(C^{-1}y+\hat{Y}_1(y,z,\mathcal{T}^{-1}(\zeta)),D^{-1}z+\hat{Z}_1(y,z,\mathcal{T}^{-1}(\zeta)),\mathcal{T}^{-1}(\zeta)).\tag{7.32}$$

Let us now seek the map H_0 in (7.21) in the prescribed form, namely

$$H_0(y,z,\zeta)=(\,y+\Lambda(y,z,\zeta)\,,\,z+\Theta(y,z,\zeta)\,),\tag{7.33}$$

where the unknowns are bounded maps Λ and Θ with values in $I\!R^e$ and $I\!R^l$ respectively. Using (7.32), one checks easily that (7.22) is equivalent to the following pair of equations:

$$\begin{aligned}\Lambda\;&=\;C\left[\hat{Y}_1+\Lambda(T_1^{-1})\right] &(7.34)\\ &\qquad+Y_2\left(C^{-1}y+\hat{Y}_1+\Lambda(T_1^{-1}),D^{-1}z+\hat{Z}_1+\Theta(T_1^{-1}),\mathcal{T}^{-1}(\zeta)\right),\\ \Theta\;&=\;D^{-1}[Z_1+\Theta(Cy+Y_1,Dz+Z_1,\mathcal{T}(\zeta))-Z_2(y+\Lambda,z+\Theta,\zeta)], &(7.35)\end{aligned}$$

where the argument of $\Lambda,\Theta,Y_i,Z_i,\hat{Y}_i,\hat{Z}_i,T_1^{-1}$, when omitted, is always (y,z,ζ). The existence of Λ and Θ will follow from another application of the shrinking lemma, this time in the space $\mathcal{B}$ of bounded functions $I\!R^e\times I\!R^l\times\mathcal{E}\to I\!R^e\times I\!R^l$ endowed with a suitable norm. More precisely, letting (Λ_1,Θ_1) denote an arbitrary member of $\mathcal{B}$ acting coordinate-wise as $(y,z,\zeta)\mapsto(\Lambda_1(y,z,\zeta),\Theta_1(y,z,\zeta))$ where Λ_1 and Θ_1 are bounded $I\!R^e$ and $I\!R^l$-valued functions respectively, we define its norm to be

$$|||(\Lambda_1,\Theta_1)|||_+=|||\Lambda_1|||+|||\Theta_1|||,$$

where $|||.|||$ indicates the *sup* norm of a map $I\!\!R^e \times I\!\!R^l \times \mathcal{E} \to I\!\!R^k$, irrespectively of k; this makes $(\mathcal{B}, |||.|||_+)$ into a a Banach space. Now, to each $(\Lambda_1, \Theta_1) \in \mathcal{B}$, we can associate another member (Λ_2, Θ_2) of $\mathcal{B}$ where $\Lambda_2 : I\!\!R^e \times I\!\!R^l \times \mathcal{E} \to I\!\!R^e$ and $\Theta_2 : I\!\!R^e \times I\!\!R^l \times \mathcal{E} \to I\!\!R^l$ are defined by

$$\begin{aligned} \Lambda_2 &= C\left[\hat{Y}_1 + \Lambda_1(T_1^{-1})\right] \qquad (7.36) \\ &\quad + Y_2\Big(C^{-1}y + \hat{Y}_1 + \Lambda_1(T_1^{-1}), D^{-1}z + \hat{Z}_1 + \Theta_1(T_1^{-1}), \mathcal{T}^{-1}(\zeta)\Big), \\ \Theta_2 &= D^{-1}\Big[Z_1 + \Theta_1(Cy + Y_1, Dz + Z_1, \mathcal{T}(\zeta)) - Z_2(y + \Lambda_1, z + \Theta_1, \zeta)\Big], \qquad (7.37) \end{aligned}$$

the argument (y, z, ζ) being omitted again for simplicity. The fact that (Λ_2, Θ_2) is indeed well-defined and belongs to $\mathcal{B}$ is a consequence of the preceding part of the proof. Consistently designating by a subscript 2 the effect of the right hand-side of (7.36) an (7.37) on some initial map, itself denoted with a subscript 1, we see from (7.20)) by inspection on (7.36) and (7.37) that, if (Λ_1, Θ_1) and (Λ_1', Θ_1') are two members of $\mathcal{B}$, then

$$\begin{aligned} |||\Lambda_2 - \Lambda_2'||| &\leq c\,|||\Lambda_1 - \Lambda_1'||| + \alpha_1|||(\Lambda_1 - \Lambda_1', \Theta_1 - \Theta_1')|||_+, \qquad (7.38) \\ |||\Theta_2 - \Theta_2'||| &\leq \frac{1}{d}\left(|||\Theta_1 - \Theta_1'||| + \alpha_1|||(\Lambda_1 - \Lambda_1', \Theta_1 - \Theta_1')|||_+\right). \qquad (7.39) \end{aligned}$$

Adding up (7.38) and (7.39), we obtain

$$\begin{aligned} &|||(\Lambda_2 - \Lambda_2', \Theta_2 - \Theta_2')|||_+ \\ &\qquad \leq [\alpha_1(1 + 1/d) + \max(c, 1/d)]\ |||(\Lambda_1 - \Lambda_1', \Theta_1 - \Theta_1')|||_+ \\ &\qquad = \alpha\,|||(\Lambda_1 - \Lambda_1', \Theta_1 - \Theta_1')|||_+ \end{aligned}$$

where by assumption $\alpha < 1$. This means that $(\Lambda_1, \Theta_1) \mapsto (\Lambda_2, \Theta_2)$ is a shrinking map on $\mathcal{B}$ whose fixed point (Λ, Θ) provides us with the unique bounded solution to (7.34) and (7.35). Equivalently, if H_0 is defined through (7.33) and R_0 through (7.21), then R_0 is the unique map $I\!\!R^e \times I\!\!R^l \times \mathcal{E} \to I\!\!R^e \times I\!\!R^l \times \mathcal{E}$ of the form (H, id), where id is the identity map on $\mathcal{E}$, such that $H - (y, z) \in \mathcal{B}$ and such that the commuting relation (7.22) holds. It remains for us to show that R_0 is a homeomorphism. For this, notice first that R_0 is continuous, because H_0 turns out to be continuous: indeed, iterating the formulas (7.36) and (7.37) starting from any initial pair (Λ_1, Θ_1) yields a sequence of maps converging to (Λ, Θ) in $\mathcal{B}$, and if the initial pair is continuous (we may for instance choose the zero map) so is every member of the sequence hence also the limit since $|||.|||_+$ induces on $\mathcal{B}$ the topology of uniform convergence. Next, if we switch the roles of T_1 and T_2, the above argument provides us with a continuous map $R_0' : I\!\!R^e \times I\!\!R^l \times \mathcal{E} \to I\!\!R^e \times I\!\!R^l \times \mathcal{E}$ of the form $(H', \mathrm{id}))$ with $H' - (y, z) \in \mathcal{B}$, satisfying $R_0'T_2 = T_1R_0'$. Then, the composed map $R = R_0'R_0$ satisfies $RT_1 = T_1R$, and since it is again of the form $(H'', \mathrm{id}))$ with $H'' - (y, z) \in \mathcal{B}$, we get $R = \mathrm{id}$ by the uniqueness part of the previous proof. Similarly $R_0R_0' = \mathrm{id}$, so that finally R_0 is invertible with continuous inverse R_0' hence a homeomorphism. $\square$

7.2 Prescribed dynamics for the control

We investigate in this subsection the situation where, in system (3.1), the control function $u(t)$ is itself the output of a dynamical system of the form:

$$\begin{array}{rcl} \dot{\zeta} & = & g(\zeta), \\ u & = & h(\zeta), \end{array} \tag{7.40}$$

where $\zeta(t) \in \mathbb{R}^q$, while $g : \mathbb{R}^q \to \mathbb{R}^q$ is locally Lipschitz continuous and $h : \mathbb{R}^q \to \mathbb{R}^m$ is continuous with, say, $h(0) = 0$. In particular, $u(t)$ is entirely determined by the finite-dimensional data $\zeta(0)$ and, from the control viewpoint, this is a particular instance of feed-forward on system (3.1) by system (7.40) where the input may only consist of Dirac delta functions.

Assume first that f is of class C^1 with respect to x and u so that (7.2) holds. Plugging (7.40) into the latter yields an ordinary differential equation in $\mathbb{R}^{n+q}$:

$$\begin{array}{rcl} \dot{x} & = & Ax + Bh(\zeta) + F(x, h(\zeta)), \\ \dot{\zeta} & = & g(\zeta). \end{array} \tag{7.41}$$

To motivate the developments to come, observe that if g is continuously differentiable with $g(0) = 0$, if A and $\partial g/\partial\zeta(0)$ are hyperbolic, and if h is continuously differentiable, then we can apply the standard Grobman-Hartman theorem on ordinary differential equations (Theorem 2.1) to conclude that the flow of (7.41) is topologically conjugate, *via* a local homeomorphism $(x, \zeta) \mapsto (z, \xi)$ around $(0, 0)$, to that of

$$\begin{pmatrix} \dot{z} \\ \dot{\xi} \end{pmatrix} = \begin{pmatrix} A & B\frac{\partial h}{\partial \zeta}(0) \\ 0 & \frac{\partial g}{\partial \zeta}(0) \end{pmatrix} \begin{pmatrix} z \\ \xi \end{pmatrix}.$$

However, the hyperbolicity requirement on $\partial g/\partial\zeta(0)$ is more stringent than it seems. Indeed, it is often desirable to study non-trivial steady behaviors, which usually entail oscillatory controls. This is why we rather seek a transformation of the form $(x, \zeta) \mapsto (H(x, \zeta), \zeta)$ that linearizes the first equation in (7.41) but preserves the second one. This can be done, as asserted by the following result which does not require hyperbolicity nor even continuous differentiability on g.

Theorem 7.4 *Suppose in system* (7.41) *that $g : \mathbb{R}^q \to \mathbb{R}^q$ is locally Lipschitz continuous, that $h : \mathbb{R}^q \to \mathbb{R}^m$ is continuous with $h(0) = 0$, that $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is continuously differentiable with $F(0, 0) = \partial F/\partial x(0, 0) = 0$, and that A is hyperbolic. Then, there exist two neighborhoods V and W of 0 in $\mathbb{R}^n$ and $\mathbb{R}^q$ respectively, and a map $H : V \times W \to \mathbb{R}^n$ with $H(0, 0) = 0$, such that*

$$\begin{array}{rccl} H \times Id : & V \times W & \to & \mathbb{R}^n \times W \\ & (x, \zeta) & \mapsto & (H(x, \zeta), \zeta) \end{array}$$

is a homeomorphism from $V \times W$ *onto its image that conjugates* (7.41) *to*

$$\begin{array}{rcl} \dot{z} & = & Az + Bh(\zeta), \\ \dot{\zeta} & = & g(\zeta). \end{array} \tag{7.42}$$

Remark 7.5 *In Theorem 7.4 (resp. Theorem 7.6 to come), we assume for convenience that all the functions involved, namely F (resp. P), g, and h, are globally defined. However, since the conclusion is local with respect to x and ζ, the same holds when these functions are only defined locally on a neighborhood of the origin, as a partition of unity argument immediately reduces the local version to the present one.*

Although it looks natural, the above theorem deserves one word of caution for the homeomorphism H depends heavily on g and h, and in a rather intricate manner. In fact, it is possible to entirely incorporate the influence of the control into the change of variables, so as to obtain a statement in which the term $Bh(\zeta)$ does not even appear in the transformed system. This will follow from Theorem 7.6 to come, for which we no longer assume in (3.1) that f is differentiable with respect to u. Accordingly, we plug (7.40) into (7.1) rather than (7.2), and we obtain instead of (7.41) the following ordinary differential equation in $\mathbb{R}^{n+q}$:

$$\begin{array}{rcl} \dot{x} & = & = Ax + P(x, h(\zeta)), \\ \dot{\zeta} & = & g(\zeta), \end{array} \tag{7.43}$$

whose flow will be denoted by $(t, x_0, \zeta_0) \mapsto (x(t, x_0, \zeta_0), \zeta(t, \zeta_0))$.

Theorem 7.6 *Suppose in system* (7.43) *that $g : \mathbb{R}^q \to \mathbb{R}^q$ is locally Lipschitz continuous, that $h : \mathbb{R}^q \to \mathbb{R}^m$ is continuous with $h(0) = 0$, that $P(x, u)$ is continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ with $P(0, 0) = 0$, that $\partial P/\partial x$ exists and is continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times n}$ with $\partial P/\partial x(0, 0) = 0$, and that A is hyperbolic. Then, there exist two neighborhoods V and W of 0 in $\mathbb{R}^n$ and $\mathbb{R}^q$ respectively, and a map $H : V \times W \to \mathbb{R}^n$ with $H(0, 0) = 0$, such that*

$$\begin{array}{rccl} H \times Id: & V \times W & \to & \mathbb{R}^n \times W \\ & (x, \zeta) & \mapsto & (H(x, \zeta), \zeta) \end{array}$$

is a homeomorphism from $V \times W$ *onto its image that conjugates* (7.43) *to*

$$\begin{array}{rcl} \dot{z} & = & Az, \\ \dot{\zeta} & = & g(\zeta), \end{array} \tag{7.44}$$

i.e. for all t, x_0, ζ_0 such that $(x(\tau, x_0, \zeta_0), \zeta(\tau, \zeta_0)) \in V \times W$ for all $\tau \in [0, t]$ (or $[t, 0]$ if $t < 0$), one has

$$H(x(t, x_0, \zeta_0), \zeta(t, \zeta_0)) = e^{tA} H(x_0, \zeta_0).$$

Theorem 7.4 is a consequence of Theorem 7.6 because the latter implies that (7.41) and (7.42) are both conjugate to (7.44). As to Theorem 7.6 itself, we will show that it is a consequence of Theorem 7.1. This will require an elementary lemma enabling us to normalize the original control system. To state the lemma, we fix, once and for all, a smooth function $\rho : [0, +\infty) \to [0, 1]$ such that

$$\left.\begin{array}{lcl} \forall t, & & |\dot\rho(t)| < 3, \\ 0 \le t \le \frac{1}{2} & \Rightarrow & \rho(t) = 1, \\ \frac{1}{2} < t < 1 & \Rightarrow & 0 < \rho(t) < 1, \\ 1 \le t & \Rightarrow & \rho(t) = 0, \end{array}\right\} \tag{7.45}$$

and we associate to any map $\beta : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ a family of functions $G_s : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, indexed by a real number $s > 0$, using the formula:

$$G_s(x,u) \triangleq \rho\left(\frac{\|x\|^2}{s^2}\right) \beta(x,u). \tag{7.46}$$

Since the context will always make clear which β is involved, our notation does not explicitly indicate the dependency of G_s on the map β. The symbol $\|.\|$, in the statement of the lemma, denotes the norm, not only of a vector, but also of a matrix; the result does not depend on a specific choice of this norm. Also, $B(x,r)$ stands for the open ball of radius r centered at x, in any Euclidean space, the open ball centered at x of radius r.

Lemma 7.7 *Let $\beta(x,u)$ be continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\partial\beta/\partial x$ continuously exist $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times n}$, with $\beta(0,0) = \partial\beta/\partial x(0,0) = 0$. Then $G_s(x,u)$ defined by (7.46) is in turn continuous and continuously differentiable with respect to x for every $s > 0$, and to each $\eta > 0$ there exist $\sigma > 0$ and $\theta > 0$ such that*

$$\forall (x,u) \in \mathbb{R}^n \times B(0,\theta)\,, \quad \|\frac{\partial G_\sigma}{\partial x}(x,u)\| \;\le\; \eta\;. \tag{7.47}$$

Proof. For the proof, we use the standard Euclidean norm on $\mathbb{R}^n$, $\mathbb{R}^m$, and the familiar operator norm on matrices. Clearly G_s is continuous and continuously differentiable with respect to x for every $s > 0$, and we have :

$$\frac{\partial G_s}{\partial x}(x,u) \;=\; \rho\left(\frac{\|x\|^2}{s^2}\right)\frac{\partial \beta}{\partial x}(x,u) \;+\; \frac{2}{s^2}\,\rho'\left(\frac{\|x\|^2}{s^2}\right)\;\beta(x,u)\;x^T, \tag{7.48}$$

where x^T is the transpose of x. Since β is continuously differentiable and $\partial\beta/\partial x\,(0,0) = 0$, we get for $s > 0$ small enough that $\|\partial\beta/\partial x\,(x,u)\| < \eta/14$ as soon as $\|x\|, \|u\| < s$. Let σ be an s with this property. Since β is continuous with $\beta(0,0) = 0$, we can in turn pick θ with $0 < \theta \le \sigma$ such that $\|\beta(0,u)\| < \eta\sigma/12$ whenever $\|u\| < \theta$. Altogether, we get that

$$\left.\begin{array}{l} \|x\| < \sigma \\ \|u\| < \theta \end{array}\right\} \Rightarrow \left\{\begin{array}{l} \|\frac{\partial\beta}{\partial x}(x,u)\| < \frac{\eta}{14}\,, \\ \|\beta(0,u)\| < \frac{\eta\sigma}{12}\,. \end{array}\right. \tag{7.49}$$

Now, we need only check (7.47) when $\|x\| < \sigma$ for otherwise G_σ is identically zero; therefore we restrict ourselves to pairs (x,u) where $\|x\| < \sigma$ and $\|u\| < \theta$. On this domain, we get from (7.49) and the mean value theorem that

$$\|\beta(x,u)\| \;\leq\; \frac{\eta}{14}\sigma \;+\; \frac{\eta\sigma}{12} \;=\; \frac{13\eta\sigma}{84}.$$

Using this together with (7.49) and the inequalities $|\rho| \leq 1$, $\|\rho'\| \leq 3$, as well as $\|x^T\| < \sigma$, formula (7.48) with $s = \sigma$ yields :

$$\|\frac{\partial G_\sigma}{\partial x}(x,u)\| \;\leq\; \frac{\eta}{14} \;+\; \frac{6}{\sigma^2}\,\frac{13\eta\sigma}{84}\,\sigma = \eta. \qquad \Box$$

Proof of Theorems 7.4 and 7.6. We already mentioned that Theorem 7.4 is a consequence of Theorem 7.6. To establish the latter, consider the following "renormalized" version of (7.43) :

$$\begin{array}{rcl} \dot{x} &=& Ax \;+\; \rho\left(\frac{\|x\|^2}{\sigma^2}\right)\; P\left(x, \rho\left(\frac{\|h(\zeta)\|}{\theta}\right) h(\zeta)\right), \\ \dot{\zeta} &=& \rho(\|\zeta\|)\, g(\zeta), \end{array} \tag{7.50}$$

where ρ is as in (7.45) and where σ, θ are strictly positive real numbers to be adjusted shortly. Because the flows of (7.50) and (7.43) do coincide as long as $\|x\| < \sigma/\sqrt{2}$, $\|\zeta\| < 1/2$, $\|h(\zeta)\| < \theta/2$, and since these inequalities define a neighborhood $(0,0)$ in $\mathbb{R}^n \times \mathbb{R}^m$ by the continuity of h and the fact that $h(0) = 0$, it is enough to prove the theorem when (7.43) *gets replaced by* (7.50) for *some pair* of strictly positive σ, θ. To this effect, we shall apply Theorem 7.1 with $\mathcal{E} = \mathbb{R}^q$ endowed with the flow of $\rho(\|\zeta\|)\, g(\zeta)$, namely $\mathcal{S}_\tau(\zeta_0)$ is the value at $t = \tau$ of the solution to the second equation in (7.50) whose value at $t = 0$ is ζ_0, and with

$$G(x,\zeta,t) = \rho\left(\frac{\|x\|^2}{\sigma^2}\right)\; P\left(x, \rho\left(\frac{\|h(\mathcal{S}_t(\zeta))\|}{\theta}\right) h(\mathcal{S}_t(\zeta))\right).$$

We now proceed to check that the assumptions of Theorem 7.1 are fulfilled if σ and θ are properly chosen. Firstly, since g is locally Lipschitz continuous while ρ is smooth with compact support on $[0,+\infty)$, we see that $\zeta \mapsto \rho(\|\zeta\|)\, g(\zeta)$ is a bounded Lipschitz continuous vector field on $\mathbb{R}^q$ hence it has a globally defined flow, which is continuous by Lemma A.1. This tells us that $(\tau,\zeta) \mapsto \mathcal{S}_\tau(\zeta)$ is continuous $\mathbb{R}\times\mathbb{R}^q \to \mathbb{R}^q$, so $\mathcal{S}_\tau$ is indeed a one-parameter group of homeomorphisms on $\mathbb{R}^q$ and $\tau \mapsto \mathcal{S}_\tau(\zeta)$ is certainly Borel measurable since it is even continuous. The continuity of $(\tau,\zeta) \mapsto \mathcal{S}_\tau(\zeta)$ also makes it clear that $G(x,\zeta,t)$ is continuous and continuously differentiable with respect to x granted the continuity of h, the smoothness of ρ, and the fact that P itself is continuous and continuously differentiable with respect to the first variable. *A fortiori* then, $x \mapsto G(x,\zeta,t)$ is continuously differentiable and $t \mapsto G(x,\zeta,t)$ is measurable.

Secondly, observe since ρ is bounded by 1 and vanishes outside $[0,1]$ that $\|\rho(\theta^{-1}\|u\|)u\| < \theta$ for all $u \in \mathbb{R}^m$, consequently G takes values in the smallest ball centered at 0 that contains $P(B(0,\sigma), B(0,\theta))$; this last set is relatively compact by the continuity of P hence G is bounded. The same argument shows that $\partial G/\partial x$ is also bounded, in other words we can choose ϕ_ζ and ψ_ζ to be suitable constant functions in (7.15), independently of ζ. In particular, (7.16) and (7.17) will hold. Moreover, if we set $\beta(x,u) = P(x,u)$, we have with the notations of (7.46) that

$$G(x,\zeta,t) = G_\sigma\left(x, \rho\left(\frac{\|h(\mathcal{S}_t(\zeta))\|}{\theta}\right) h(\mathcal{S}_t(\zeta))\right). \tag{7.51}$$

Since $\rho(\theta^{-1}\|h(v)\|)h(v)$ lies in $B(0,\theta)$ for all $v \in I\!R^q$ so in particular for $v = \mathcal{S}_t(\zeta)$, we deduce from (7.51) and Lemma 7.7 that $\partial G/\partial x$ can be made uniformly small for suitable σ and θ. That is to say the number η in (7.17) can be made arbitrarily small upon choosing σ and θ adequately, in particular we can meet (7.18).

Thirdly, the condition (7.15) that we just proved to hold (actually with constant functions ϕ_ζ and ψ_ζ independent of ζ) entails that the first equation in (7.50) has a unique solution given initial conditions $x(0)$ and $\zeta(0)$ (*cf* for instance [22, Theorem 54, Proposition C.3.4, Proposition C.3.8]) and, since the same holds true for the second equation as was pointed out when we defined $\mathcal{S}_\tau(\zeta)$, we conclude that the whole vector field in the right hand-side of (7.15) has a flow on $I\!R^{n+q} = I\!R^n \times I\!R^q$, which is continuous by Lemma A.1. As $\widehat{x}$, defined in (7.5), is nothing but the projection of this flow onto the first factor $I\!R^n$, we conclude that $(\tau, x_0, \zeta) \mapsto \widehat{x}(\tau, x_0, \zeta)$ is continuous. Finally, notice that (7.4) is immediate from the group property of $\mathcal{S}_\tau$. Having verified all the hypotheses of Theorem 7.1, we apply the latter to conclude the proof of Theorem 7.6. $\square$

7.3 Control systems viewed as flows

In [5], a general way of associating a flow to a control system is proposed, based on the action of the time shift on some functional space of inputs. Before giving the proper framework for our results, let us first carry out a few measure-theoretic preliminaries.

For arbitrary exponents $p \in [1,\infty]$, we denote by $\mathcal{L}^p(I\!R, I\!R^m)$, or simply by $\mathcal{L}^p$ for short, the space of measurable functions $\Upsilon : I\!R \to I\!R^m$ such that

$$\begin{aligned} \|\Upsilon\|_p &= \left(\int_{I\!R} \|\Upsilon(t)\|^p\, dt\right)^{1/p} < \infty \qquad \text{if } p < \infty, \\ \|\Upsilon\|_\infty &= \operatorname*{ess.\,sup.}_{t \in I\!R} \|\Upsilon(t)\| \quad < \infty \qquad \text{if } p = \infty. \end{aligned}$$

In the above, measurability and summability were implicitly understood with respect to Lebesgue measure. However, the same definitions can of course be made for any positive measure. We only consider measures defined on the same σ-algebra as Lebesgue measure (namely the completion of the Borel σ-algebra with respect to sets of Lebesgue measure zero). We explicitly indicate the dependence on the measure μ of the corresponding functional spaces and norms by writing $\mathcal{L}^{p,\mu}$ and $\|.\|_{p,\mu}$.

Remark 7.8 *If μ is a positive measure on $I\!R$ as described above, and if μ and Lebesgue measure are mutually absolutely continuous, then for any Lebesgue measurable (hence also*

μ-measurable) function Υ it holds that $\|\Upsilon\|_\infty = \|\Upsilon\|_{\infty,\mu}$. Indeed, we have that $\|\Upsilon\|_\infty \leq \alpha$ if, and only if, the set E_α of those $x \in \mathbb{R}$ for which $\|\Upsilon\|(x) > \alpha$ has Lebesgue measure zero. Since the latter holds if, and only if, $\mu(E_\alpha) = 0$, it is equivalent to require that $\|\Upsilon\|_{\infty,\mu} \leq \alpha$ as announced.

For any $p \in [1, \infty]$ and $\tau \in \mathbb{R}$, we define the time shift $\Theta_\tau : \mathcal{L}^p \to \mathcal{L}^p$ by

$$\Theta_\tau(\Upsilon)(t) = \Upsilon(\tau + t) \ . \tag{7.52}$$

It is well known that, for fixed $\Upsilon \in \mathcal{L}^p$, the map $\tau \mapsto \Theta_\tau(\Upsilon)$ is continuous $\mathbb{R} \to \mathcal{L}^p$ if $1 \leq p < \infty$ [21, Theorem 9.5]. When $p = \infty$ it is no longer so, but the map is at least Borel measurable :

Lemma 7.9 *For fixed $\Upsilon \in \mathcal{L}^\infty$, consider the map $T_\Upsilon : \mathbb{R} \to \mathcal{L}^\infty$ defined by $T_\Upsilon(\tau) = \Theta_\tau(\Upsilon)$. If V is open in $\mathcal{L}^p$, then $T_\Upsilon^{-1}(V)$ is measurable in $\mathbb{R}$.*

Proof. Set for simplicity $T_\Upsilon(\tau) = \Upsilon_\tau$, and fix arbitrarily $v \in \mathcal{L}^\infty$ together with $\varepsilon > 0$. It is enough to show that the set

$$E = \{\tau \in \mathbb{R}; \ \ \|\Upsilon_\tau - v\|_\infty > \varepsilon\}$$

is measurable. Let μ be the measure on $\mathbb{R}$ such that $d\mu(t) = dt/(1+t^2)$. In view of Remark 7.8, we can replace $\|.\|_\infty$ by $\|.\|_{\infty,\mu}$ in the definition of E. Now, since μ is finite, the functions Υ_τ and v belong to $\mathcal{L}^{1,\mu}$, which is to the effect that

$$\lim_{p\to\infty} \|\Upsilon_\tau - v\|_{p,\mu} = \|\Upsilon_\tau - v\|_{\infty,\mu}, \tag{7.53}$$

see *e.g.* [21, Chap. 3, Ex.4]. In particular, if we let

$$E_{p,\mu} = \{\tau \in \mathbb{R}; \ \ \|\Upsilon_\tau - v\|_{p,\mu} > \varepsilon\},$$

we deduce from (7.53) that

$$E = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} E_{j,\mu}$$

where k and j assume integral values, so we are left to prove that $E_{j,\mu}$ is measurable. But since translating the argument is a continuous operation $\mathbb{R} \to \mathcal{L}^{p,\mu}$ when $p < \infty$ [21, Theorem 9.5][4], each $E_{j,\mu}$ is in fact open in $\mathbb{R}$ thereby proving the lemma. $\square$

Endowed with $\|.\|_p$-balls as neighborhoods of 0, the set $\mathcal{L}^p$ is a topological vector space but it is not Hausdorff; identifying functions that agree almost everywhere, we obtain the familiar Lebesgue space L^p of equivalence classes of $\mathcal{L}^p$-functions; it is a Banach space,

[4]The proof is given there for Lebesgue measure only, but it does carry over *mutatis mutandis* to any complete regular Borel measure on $\mathbb{R}$, hence in particular to μ.

whose norm, still denoted by $\|.\|_p$, is induced by $\|.\|_p$ defined in $\mathcal{L}^p$, and whose topology coincides with the quotient topology arising from the canonical map $\mathcal{L}^p \to L^p$. The time shift $\Theta_\tau : \mathcal{L}^p \to \mathcal{L}^p$ defined by (7.52) induces a well defined map $\Theta_\tau : L^p \to L^p$. In what follows, results are stated in terms of L^p, but we do make use of $\mathcal{L}^p$ for the proof because point-wise evaluation makes no sense in L^p.

Let us now come back to our control system, namely (7.1), which is obtained from (3.1) by singling out the linear term in x around the equilibrium $(0,0) \in \mathbb{R}^n \times \mathbb{R}^m$. This time, however, we emphasize the functional dependence on the control by writing

$$\dot{x} \;=\; Ax \;+\; P(x, \Upsilon(t)), \tag{7.54}$$

where, as in the preceding subsection, $P : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is continuous and has continuous derivative with respect to the first argument $\frac{\partial P}{\partial x} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times n}$. Also, we fix some $p \in [1, \infty]$ will consider controls $\Upsilon \in L^p(\mathbb{R}, \mathbb{R}^m)$. Note that, unless $p = \infty$, we are departing ourselves from the convention made in Definition 3.1 that inputs are locally bounded. When $p < \infty$, to take this fact into account, we assume moreover that, to each compact set $K \subset \mathbb{R}^n$, there are positive constants $c_1(K)$, $c_2(K)$ such that

$$\|P(x,u)\| + \|\frac{\partial P}{\partial x}(x,u)\| \le c_1(K) + c_2(K)\,\|u\|^p, \quad (x,u) \in K \times \mathbb{R}^m, \tag{7.55}$$

where we agree, for definiteness, that the norm of a matrix is the operator norm. Classical results imply (see *e.g.* [22, Theorem 54, Proposition C.3.4]) that the solution to (7.54) uniquely exists on some maximal time interval once $x(0) = x_0$ and $\Upsilon \in L^p$ are chosen. This solution we denote by

$$t \mapsto x(t, x_0, \Upsilon)\ .$$

This allows one to define a flow on $\mathbb{R}^n \times \mathcal{L}^p$, or on $\mathbb{R}^n \times L^p$, the flow at time τ being given by

$$(x_0, \Upsilon) \;\mapsto\; (\,x(\tau, x_0, \Upsilon)\,,\, \Theta_\tau(\Upsilon)\,)\ . \tag{7.56}$$

The main result in this subsection is the following. It is of purely *open loop* character, that is to say the linearizing transformation $(x, \Upsilon) \mapsto (z, \Upsilon)$ operates at a functional level where z depends not only on x, but also on the whole input function $\Upsilon : \mathbb{R} \mapsto \mathbb{R}^m$. That type of linearization is intriguing in the authors' opinion, but its usefulness in control is not clear unless the structure of the transformation is thoroughly understood. Unfortunately our method of proof does not reveal much in this direction, which may deserve further study.

Theorem 7.10 *Suppose in* (7.54) *that $P(x,u)$ is continuous $\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^n$ with $P(0,0) = 0$, that $\partial P/\partial x$ exists and is continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times n}$ with $\partial P/\partial x(0,0) = 0$, and that A is hyperbolic. Let $p \in [1,\infty]$, and,if $p < \infty$, assume that, to each compact set $K \subset \mathbb{R}^n$, there are positive constants $c_1(K)$, $c_2(K)$ such that* (7.55) *holds. Then, there exist two neighborhoods V and W of 0 in $\mathbb{R}^n$ and $L^p(\mathbb{R}, \mathbb{R}^m)$ respectively, and a map $H : V \times W \to \mathbb{R}^n$ with $H(0,0) = 0$, such that*

$$\begin{array}{rccc} H\times Id: & V\times W & \to & \mathbb{R}^n\times W \\ & (x,\Upsilon) & \mapsto & (H(x,\Upsilon),\Upsilon) \end{array} \tag{7.57}$$

is a homeomorphism from $V \times W$ onto its image that conjugates (7.54) to

$$\dot{z} = Az, \tag{7.58}$$

i.e. for all $(t, x_0, \Upsilon) \in \mathbb{R} \times \mathbb{R}^n \times L^p(\mathbb{R}, \mathbb{R}^m)$ such that $(x(\tau, x_0, \Upsilon), \Upsilon) \in V \times W$ for all $\tau \in [0, t]$ (or $[t, 0]$ if $t < 0$) one has

$$H(x(t, x_0, \Upsilon)) = e^{tA} H(x_0, \Upsilon). \tag{7.59}$$

Remark 7.11 *The above theorem parallels Theorem 7.6 of section 7.2, in that we initially wrote $\dot{x} = f(x, u)$ in the form (7.1), assuming that f is continuously differentiable with respect to x, to finally conclude, under suitable hypotheses, that (7.54) is locally conjugate in some appropriate sense to the non-controlled linear system (7.58). We might as well have stated an analog to Theorem 7.4 where, assuming this time that f is of class C^1, we write $\dot{x} = f(x, u)$ in the form (7.2) with hyperbolic A, assuming in addition if $p < \infty$ that for any compact $K \subset \mathbb{R}^n$ one has*

$$\|F(x, u)\| + \|\frac{\partial F}{\partial x}(x, u)\| \leq c_1(K) + c_2(K)\, \|u\|^p, \quad (x, u) \in K \times \mathbb{R}^m, \tag{7.60}$$

to conclude that $\dot{x} = Ax + B\Upsilon(t) + F(x, \Upsilon(t))$ is conjugate via $z = H(x, \Upsilon)$ to $\dot{z} = Az + B\Upsilon(t)$, where $H \times Id$ is a local homeomorphism at 0×0 of $\mathbb{R}^n \times L^p$. Again, although the presence of the control term $B\Upsilon(t)$ in the linearized equation makes it look more natural, the result we just sketched is a logical consequence of Theorem 7.10 just like Theorem 7.4 was a consequence of Theorem 7.6.

To prove Theorem 7.10 we shall again apply Theorem 7.1 to a suitably normalized version of (7.54), the normalization step depending on the following lemma which stands analogous to Lemma 7.7 in the $\mathcal{L}^p$ context. For convenience, we denote below by $B_{\mathcal{L}^p}(v, r)$ the ball centered at v of radius r in $\mathcal{L}^p$, and by $\mathcal{L}^1_{loc}(\mathbb{R}, \mathbb{R}^m)$ (or simply $\mathcal{L}^1_{loc}$ if no confusion can arise) the space of *locally integrable functions*, namely those whose restriction to any compact $K \subset \mathbb{R}$ belongs to $\mathcal{L}^1(K, \mathbb{R}^m)$.

Lemma 7.12 *Let $\beta(x, u)$ be continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $\partial\beta/\partial x$ continuously exist $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{n\times n}$, with $\beta(0, 0) = \partial\beta/\partial x(0, 0) = 0$. Assume for some $p \in [1, \infty)$ that, to each compact set $K \subset \mathbb{R}^n$, there are positive constants $c_1(K)$, $c_2(K)$ such that*

$$\|\beta(x, u)\| + \|\frac{\partial \beta}{\partial x}(x, u)\| \leq c_1(K) + c_2(K)\, \|u\|^p, \quad (x, u) \in K \times \mathbb{R}^m. \tag{7.61}$$

Then, G_s being as in (7.46), it holds for every $s > 0$ and any $\Upsilon \in \mathcal{L}^p(\mathbb{R}, \mathbb{R}^m)$ that $G_s(x, \Upsilon) \in \mathcal{L}^1_{loc}(\mathbb{R}, \mathbb{R}^n)$ and $\partial G_s/\partial x(x, \Upsilon) \in \mathcal{L}^1_{loc}(\mathbb{R}, \mathbb{R}^{n\times n})$ for fixed $x \in \mathbb{R}$. Moreover, to each $\eta > 0$ there exist $\sigma > 0$ and $\theta > 0$ such that G_σ satisfies :

$$\begin{array}{lll} \forall \Upsilon \in B_{\mathcal{L}^p}(0, \theta)\,, & \text{there exists} \;\; \psi_\Upsilon \in \mathcal{L}^1_{loc}(\mathbb{R}, \mathbb{R}) \;\; \text{such that} & \\ \|\psi_\Upsilon\|_{L^1[0,1]} \leq \eta & \text{and,} \;\; \forall x \in \mathbb{R}^n\,, \;\; \|\frac{\partial G_\sigma}{\partial x}(x, \Upsilon)\| \leq \; \psi_\Upsilon. \end{array} \tag{7.62}$$

Proof. For fixed $x \in I\!R$, it is clear from (7.61) that $G_s(x,\Upsilon)$ and $\partial G_s/\partial x(x,\Upsilon)$ belong to $\mathcal{L}^1_{loc}(I\!R, I\!R^n)$ when $\Upsilon \in \mathcal{L}^p(I\!R, I\!R^m)$, measurability being ensured by the continuity of G_s and $\partial G_s/\partial x$. To prove (7.62), first apply Lemma 7.7 to find $\sigma > 0$ and $\theta_0 > 0$ such that

$$\forall (x,u) \in I\!R^n \times B(0,\theta_0)\,, \quad \|\frac{\partial G_\sigma}{\partial x}(x,u)\| \;\leq\; \eta/2\;. \tag{7.63}$$

Next, let $c_1 = c_1(\overline{B}(0,\sigma))$ and $c_2 = c_2(\overline{B}(0,\sigma))$ be defined after (7.61), and observe that

$$\forall (x,u) \in I\!R^n \times I\!R^m, \quad \|\partial G_\sigma/\partial x\,(x,u)\| \leq (1+6/\sigma)(c_1 + c_2\|u\|^p) \tag{7.64}$$

because when $\|x\| < \sigma$ this follows from (7.48), (7.61) and the fact that $|\rho'| < 3$, whereas G_σ vanishes anyway when $\|x\| \geq \sigma$. Introduce now the set

$$E_{\Upsilon,\theta_0} = \{t \in [0,1], \quad \|\Upsilon\| < \theta_0\}. \tag{7.65}$$

Letting $\psi_\Upsilon(t) = \eta/2$ for $t \in E_{\Upsilon,\theta_0}$ and $\psi_\Upsilon(t) = (1+6/\sigma)(c_1+c_2\|\Upsilon(t)\|^p)$ otherwise, it is clear that $\psi_\Upsilon \in \mathcal{L}^1_{loc}(I\!R, I\!R)$ and it follows from (7.65), (7.63), and (7.64) that $\|\partial G_{\sigma_0}/\partial x(x,\Upsilon)\| \leq \psi_\Upsilon$ for any $x \in I\!R^n$. In another connection, let ν be the measure on $I\!R$ given by $d\nu(t) = |\Upsilon(t)|^p dt$. By absolute continuity of ν with respect to Lebesgue measure, there is $\varepsilon > 0$ such that

$$\int_E \|\Upsilon\|^p\,dt < \frac{\eta}{4c_2(1+6/\sigma)} \qquad \text{as soon as } |E| < \varepsilon, \tag{7.66}$$

where $|E|$ denotes the Lebesgue measure of a measurable set $E \subset I\!R$ [21, Theorem 6.11]. Pick $\theta > 0$ so small that

$$\frac{\theta}{\theta_0} < \max\left\{\varepsilon, \frac{\eta}{4(1+6/\sigma)c_1}\right\}. \tag{7.67}$$

Then, if $\|\Upsilon\|_p < \theta$, the set $[0,1]\setminus E_{\Upsilon,\theta_0}$ has measure at most θ/θ_0 hence, by definition of ψ_Υ, we get in view of (7.66) and (7.67) the estimate :

$$\|\psi_\Upsilon\|_{L^1[0,1]} \leq \frac{\eta}{2} + \frac{\theta}{\theta_0}(1+6/\sigma)c_1 + \frac{\eta}{4}$$

which is less that $\eta/2 + \eta/4 + \eta/4 = \eta$ by (7.67) again, as desired. $\square$

We are now in position to establish Theorem 7.10.

Proof of Theorem 7.10 For the proof we can replace L^p by $\mathcal{L}^p$, because if we find a local homeomorphism of $I\!R^n \times \mathcal{L}^p$ at 0×0, of the form $\widetilde{H}\times Id$, that conjugates (7.54) to (7.58), the fact that $x(\tau, x_0, \Upsilon)$ depends only on the equivalence class of Υ in L^p implies that the same holds true for $\widetilde{H}(x_0,\Upsilon)$, and therefore $\widetilde{H}\times Id$ will induce a quotient map $H \times Id$ around 0×0 in $I\!R^n \times L^p$ that is still a local homeomorphism by definition of the quotient topology. To prove the $\mathcal{L}^p$ version, we consider the following "re-normalization" of (7.54) :

$$\dot{x} \;=\; Ax \;+\; \rho\left(\frac{\|x\|^2}{\sigma^2}\right)\; P\left(x, \rho\left(\frac{\|\Upsilon\|_p}{\theta}\right)\Upsilon\right), \tag{7.68}$$

where ρ is as in (7.45) and σ, θ are strictly positive real numbers to be fixed. Because the right-hand sides of (7.68) and (7.54) agree as long as $\|x\| < \sigma/\sqrt{2}$ and $\|\Upsilon\|_p < \theta/2$ which defines a neighborhood $(0,0)$ in $\mathbb{R}^n \times \mathcal{L}^p$, it is enough to prove the theorem when (7.54) *gets replaced by* (7.68) for *some pair* σ, θ. To this effect, we shall apply Theorem 7.1 with $\mathcal{E} = \mathcal{L}^p$, endowed with the one-parameter group of transformations $\mathcal{S}_\tau = \Theta_\tau$ defined by (7.52), and

$$G(x,\zeta,t) = \rho\left(\frac{\|x\|^2}{\sigma^2}\right)\ P\left(x, \rho\left(\frac{\|\zeta\|_p}{\theta}\right)\zeta(t)\right).$$

Let us check that the assumptions of Theorem 7.1 are met if σ and θ are suitably chosen.

Firstly, it is obvious that $\mathcal{S}_\tau$ is continuous (hence a homeomorphism since $\mathcal{S}_\tau^{-1} = \mathcal{S}_{-\tau}$) because it is a linear isometry of $\mathcal{L}^p$. In addition, $\tau \mapsto \mathcal{S}_\tau(\zeta)$ is certainly Borel measurable, because it is even continuous when $p < \infty$ [21, Theorem 9.5] while Lemma 7.9 applies if $p = \infty$.

Secondly, it follows immediately from the assumptions on P and the smoothness of ρ that $G(x,\zeta,t)$ is continuously differentiable with respect to x for fixed ζ and t, while the measurability of $t \mapsto G(x,\zeta,t)$ follows from the continuity of P and the measurability of ζ. To prove the existence of ϕ_ζ and ψ_ζ in (7.15), we distinguish between $p < \infty$ and $p = \infty$. If $p < \infty$, by (7.55) and the fact that ρ is bounded by 1 and vanishes outside $[0,1]$, a valid choice for ϕ_ζ is

$$\phi_\zeta(t) = c_1(\,\overline{B}(0,\sigma)\,) + c_2(\,\overline{B}(0,\sigma)\,)\ \rho^p\left(\frac{\|\zeta\|_p}{\theta}\right)\|\zeta(t)\|^p$$

and, since by the properties of ρ we have that

$$\left\|\rho\left(\frac{\|\zeta\|_p}{\theta}\right)\zeta\right\|_p \leq \theta \quad \forall \zeta \in \mathcal{L}^p,\ \ 1 \leq p \leq \infty, \tag{7.69}$$

it follows that (7.16) is met with

$$M = c_1(\,\overline{B}(0,\sigma)\,) + c_2(\,\overline{B}(0,\sigma)\,)\ \theta^p.$$

As to ψ_ζ, observe if we set $\beta(x,u) = P(x,u)$ that, with the notations of (7.46), one has

$$G(x,\zeta,t) = G_\sigma\left(x, \rho\left(\frac{\|\zeta\|_p}{\theta}\right)\zeta(t)\right), \tag{7.70}$$

so Lemma 7.12 ensures the existence of ψ_ζ and also that the number η in (7.17) can be made arbitrarily small upon choosing σ and θ adequately; in particular we can meet (7.18). If $p = \infty$, we let

$$\phi_\zeta(t) = \sup_{x \in \overline{B}(0,\sigma)} \left\|P\left(x, \rho\left(\frac{\|\zeta\|_p}{\theta}\right)\zeta(t)\right)\right\|$$

so that the first half of (7.15) holds by the properties of ρ. By (7.69) we also have that

$$\|\phi_\zeta\|_\infty \leq \sup_{(x,u)\in\overline{B}(0,\sigma)\times\overline{B}(0,\theta)} \|P(x,u)\|, \tag{7.71}$$

so that $\phi_\zeta \in \mathcal{L}^\infty(I\!R, I\!R)$ hence it is locally summable, and the right-hand side of (7.71) may serve as M in (7.16). As to ψ_ζ, observe that (7.70) still holds for $p = \infty$, again with $\beta(x,u) = P(x,u)$, so we can set

$$\psi_\zeta(t) = \sup_{x\in\overline{B}(0,\sigma)} \left\| \frac{\partial G_\sigma}{\partial x} \left(x, \rho\left(\frac{\|\zeta\|_\infty}{\theta}\right)\zeta(t)\right)\right\|,$$

and using (7.69) once more we get

$$\|\psi_\zeta\|_\infty \leq \sup_{(x,u)\in\overline{B}(0,\sigma)\times\overline{B}(0,\theta)} \left\| \frac{G_\sigma}{\partial x}(x,u)\right\|. \tag{7.72}$$

Thus $\psi_\zeta \in \mathcal{L}^\infty(I\!R, I\!R)$ hence it is locally summable, and applying Lemma 7.7 to the right-hand side of (7.72) shows that $\|\psi_\zeta\|_\infty$ can be made arbitrarily small upon choosing σ and θ adequately. Consequently η in (7.17) can be as small as we wish and in particular we can meet (7.18).

Thirdly, $t \mapsto \widehat{x}(t, x_0, \zeta)$ defined in (7.5) is just the solution to (7.68) corresponding to $\Upsilon = \zeta$ and $x(0) = x_0$, which uniquely exists for all t by (7.15), see *e.g.* [22, Theorem 54, Proposition C.3.4, Proposition C.3.8]. The continuity $I\!R^n \times \mathcal{L}^p \to I\!R^n$ of $(x_0, \zeta) \mapsto \widehat{x}(t, x_0, \zeta)$ is now ascertained by Proposition C.1, once it is observed that $F(x,u) = Ax + \rho(\|x\|^2/\sigma^2)P(x,u)$ satisfies the hypotheses of that proposition by (7.55) and the properties of ρ, and that $Ax + G(x, \zeta, t)$ is just the composition of F with the continuous map on $I\!R^n \times \mathcal{L}^p$ given by $(x, \zeta) \mapsto (x, \rho(\|\zeta\|_p/\theta)\zeta)$ (Proposition C.1 was actually proved for L^p controls, but nothing is to be changed if we work in $\mathcal{L}^p$).

Finally, notice that (7.4) is immediate by the very definition of Θ_τ. Thus we can apply Theorem 7.1 to conclude the proof of Theorem 7.10. $\square$

Remark 7.13 *It should be noted that, unlike Theorems 7.4 and 7.6, Theorem 7.10 cannot be localized with respect to u when $p < \infty$. However, using a partition of unity argument, the result carries over to the case where, in* (7.54)*, the map P is only defined on $\mathcal{V} \times I\!R^m$ where $\mathcal{V}$ is a neighborhood of* 0 *in $I\!R^n$.*

In [5], particular attention is payed to the weak-* topology on $\mathcal{L}^\infty$ for the control space, because it makes the flow $\tau \mapsto \Theta_\tau(\Upsilon)$ continuous for fixed Υ. Subsequently, this reference focuses on systems that are affine in the control : $\dot{x} = X_0(x) + C(x)u$, where X_0 is a C^1 vector field on $I\!R^n$ and $C : I\!R^n \to I\!R^{n\times m}$ a C^1 matrix-valued function; the reason for this affine restriction is that it ensures, in the weak-* context, the sequential continuity of $(x_0, \Upsilon) \mapsto x(\tau, x_0, \Upsilon)$ for fixed τ, whenever the flow makes sense : this is easily deduced from the Ascoli-Arzela theorem and the fact that weak-* convergent sequences are norm-bounded [20, Theorem 2.5]. Although the continuity of the flow Θ was never a concern to us (only Borel measurability was required), it is natural in this connection to ask what happens with Theorem 7.10 if we endow L^∞ with the weak-* topology inherited from the (L^1, L^∞) duality. On the one hand, *in case one restricts his attention, as is done in [5], to a balanced,* weak-* compact *time-shift invariant subset of* L^∞ *containing* 0, *e.g. a ball* $\bar{B}_{L^\infty}(0,r)$, *then the conclusions of the theorem still hold if we equip the subset in question with the weak-* topology.* Indeed, the weak-* topology is metrizable on any compact set E because L^1 is separable [20, Theorems 3.16] and, since weak-* convergent sequences are norm- bounded, it follows if E is balanced that one can find a neighborhood of 0 in E which is included in $\bar{B}_{L^\infty}(0,\theta)$ for arbitrary small θ. In particular we can embed this neighborhood in W of Theorem 7.10, and then it only remains to show that (7.57) remains continuous if W is equipped with the weak-* topology; this in turn reduces *via* (7.59) to the already mentioned fact that $(x_0, \Upsilon) \mapsto x(\tau, x_0, \Upsilon)$ is sequentially continuous for fixed τ when the topology on Υ is the weak-* one. On the other hand, working weak-* with unrestricted controls in L^∞ raises serious difficulties, for no weak-* neighborhood in L^∞ can be norm-bounded. This results in the fact that, although Θ is now continuous, the domain of definition of the flow (7.56) may fail to be open : for instance the equation $\dot{x} = x + x^2\Upsilon(t)$ with initial condition $x(0) = x_0$, where x and Υ are real-valued, cannot have a solution on a fixed interval $[0, t]$ for every $(x_0, \Upsilon) \in B(0,r) \times \mathcal{W}_0$ if $\mathcal{W}_0$ is a weak-* neighborhood of 0 in $L^\infty(\mathbb{R}, \mathbb{R})$. Therefore it is hopeless to build a local homeomorphism by integrating the flow as is done in the proof of Theorem 7.1, and the authors do not know what analog to Theorem 7.10 could be carried out in this context.

Remark 7.14 *We mentioned in section 3.2 the paper [3], where transformations* $\mathbb{R}^n \times L^\infty \to \mathbb{R}^n \times L^\infty$ *are also considered, using for the input space a topology on* L^∞ *which is intermediate between the weak-* and the strong one. There the structure of conjugating homeomorphisms is not* (7.57) *but rather a triangular form reminiscent of* (3.10) *:*

$$(x, \Upsilon) \quad \mapsto \quad (\,H(x)\,,\,F(x, \Upsilon)\,)$$

that combines what is called in this reference "topological static state feedback equivalence" and "topological state equivalence"[3, Definition 5]. We refer the interested reader to the original paper for a result on topological linearization of systems with two states and one control, using this type of transformation, under some global hypotheses.

APPENDIX (A to F)

A Three basic lemmas on ODEs

Throughout this section, we let $\mathcal{U}$ be an open subset of $\mathbb{R}^d$. We say that a continuous vector field $X : \mathcal{U} \to \mathbb{R}^d$ *has a flow* if the Cauchy problem $\dot{x}(t) = X(x(t))$ with initial condition $x(0) = x_0$ has a unique solution, defined for $t \in (-\varepsilon, \varepsilon)$ with $\varepsilon = \varepsilon(x_0) > 0$. The flow of X at time t is denoted by X_t, in other words we have with the preceding notations that $X_t(x_0) = x(t)$. It is easy to see that the domain of definition of $(t, x) \mapsto X(t, x)$ is open in $\mathbb{R} \times \mathcal{U}$.

Lemma A.1 *If $X : \mathcal{U} \to \mathbb{R}^d$ is a continuous vector field that has a flow, the map $(t, x) \mapsto X_t(x)$ is continuous on the open subset of $\mathbb{R} \times \mathcal{U}$ where it is defined.*

Proof. This is an easy consequence of the Ascoli-Arzela theorem, and actually a special case of [8, chap. V, Theorem 2.1]. □

Lemma A.2 *Assume that the sequence of continuous vector fields $X^k : \mathcal{U} \to \mathbb{R}^d$ converges to X, uniformly on compact subsets of $\mathcal{U}$, and that all the X^k as well as X itself have a flow. Suppose that $X_t(x)$ is defined for all $(t, x) \in [0, T] \times K$ with $T > 0$ and $K \subset \mathcal{U}$ compact. Then $X_t^k(x)$ is also defined on $[0, T] \times K$ for k large enough, and the sequence of mappings $(t, x) \mapsto X_t^k(x)$ converges to $(t, x) \mapsto X_t(x)$, uniformly on $[0, T] \times K$.*

Proof. By assumption,

$$K_1 = \{X_t(x);\ (t, x) \in [0, T] \times K\}$$

is a well-defined subset of $\mathcal{U}$ that contains K, and it is compact by Lemma A.1. Let K_0 be another compact subset of $\mathcal{U}$ whose interior contains K_1, and put $d(K_1, \mathcal{U} \setminus K_0) = \eta > 0$ where $d(E_1, E_2)$ indicates the distance between two sets E_1, E_2. From the hypothesis there is $M > 0$ such that $\|X^k\| \leq M$ on K_0 for all k, hence the maximal solution to $\dot{x}(t) = X^k(x(t))$ with initial condition $x(0) = x_0 \in K$ remains in K_0 as long as $t \leq \eta/2M$. Consequently the flow $(t, x) \mapsto X_t^k(x)$ is defined on $[0, \eta/2M] \times K$ for all k, with values in K_0. *We claim* that it is a bounded equicontinuous sequence of functions there. Boundedness is clear since these functions are K_0-valued, so we must show that, to every $(t, x) \in [0, \eta/2M] \times K$ and every $\varepsilon > 0$, there is $\alpha > 0$ such that $\|X^k(t', x') - X^k(t, x)\| < \varepsilon$ for all k as soon as $|t - t'| + \|x - x'\| < \alpha$. By the mean-value theorem and the uniform majorization $\|X^k(X_t^k(x))\| \leq M$, it is sufficient to prove this when $t = t'$. Arguing by contradiction, assume for some subsequence k_l and some sequence x_l converging to x in K that

$$\|X_t^{k_l}(x) - X_t^{k_l}(x_l)\| \geq \varepsilon \quad \text{for all } l \in \mathbb{N}. \tag{A.1}$$

Then, by Lemma A.1, the index k_l tends to infinity with l. Next consider the sequence of maps $F_l : [0, \eta/2M] \to K_0$ defined by $F_l(t) = X_t^{k_l}(x_l)$. Again, by the mean value theorem, it is a bounded equicontinuous family of functions and, by the Ascoli-Arzela theorem, it is relatively compact in the topology of uniform convergence (compare [8, chap. II, Theorem

3.2]). But if $\Phi : [0, \eta/2M] \to K_0$ is the uniform limit of some subsequence F_{l_j}, and since $X^{k_{l_j}}$ converges uniformly to X on K_0 as $j \to \infty$, taking limits in the relation

$$X_t^{k_{l_j}}(x_{l_j}) = x_{l_j} + \int_0^t X^{k_{l_j}}(X_s^{k_{l_j}}(x_{l_j}))\, ds$$

gives us

$$\Phi(t) = x + \int_0^t X(\Phi(s))\, ds$$

so that $\Phi(t) = X_t(x)$ since X has a flow. Altogether $F_l(t)$ converges uniformly to $X_t(x)$ on $[0, \eta/2M]$ because this is the only accumulation point, and then (A.1) becomes absurd. *This proves the claim.* From the claim it follows, using the Ascoli-Arzela theorem again, that the family of functions $(t, x) \mapsto X_t^k(x)$ is relatively compact for the topology of uniform convergence $[0, \eta/2M] \times K \to K_0$, and in fact it converges to $(t, x) \mapsto X_t(x)$ because, by the same limiting argument as was used to prove the claim, every accumulation point $\Phi(t, x)$ must be a solution to

$$\Phi(t, x) = x + \int_0^t X(s, \Phi(s, x))\, ds$$

hence for fixed x is an integral curve of X with initial condition x. In particular, by definition of K_1, we shall have that $d(X_t^k(x), K_1) < \eta/2$ for all $(t, x) \in [0, \eta/2M] \times K$ as soon as k is large enough. For such k the flow $(t, x) \mapsto X_t^k(x)$ will be defined on $[0, \eta/M] \times K$ with values in K_0, and we can repeat the whole argument again to the effect that $X_t^k(x)$ converges uniformly to $X_t(x)$ there. Proceeding inductively, we obtain after $[2TM/\eta] + 1$ steps at most that $(t, x) \mapsto X_t^k(x)$ is defined on $[0, T] \times K$ with values in K_0 for k large enough, and converges uniformly to $(t, x) \mapsto X_t(x)$ there, as was to be shown. $\square$

The next lemma stands analogous to Lemma A.2 for time-dependent vector fields, assuming that the convergence holds boundedly almost everywhere in time. The assumption that the vector fields have a flow is replaced here by a local Lipschitz condition that we now comment upon.

By definition, a time-dependent vector field $X : [t_1, t_2] \times \mathcal{U} \to \mathbb{R}^d$ is locally Lipschitz with respect to the second variable if every $(t_0, x_0) \in [t_1, t_2] \times \mathcal{U}$ has a neighborhood there such that $\|X(t, x') - X(t, x)\| < c\|x' - x\|$, for some constant c, whenever (t, x) and (t, x') belong to that neighborhood. This of course entails that X is bounded on compact subsets of $[t_1, t_2] \times \mathcal{U}$. Next, by the compactness of $[t_1, t_2]$, the local Lipschitz character of X strengthens to the effect that each $x_0 \in \mathcal{U}$ has a neighborhood $\mathcal{N}_{x_0}$ such that $\|X(t, x') - X(t, x)\| < c_{x_0}\|x' - x\|$, for some constant c_{x_0}, whenever $x, x' \in \mathcal{N}_{x_0}$ and $t \in [t_1, t_2]$. If now $\mathcal{K} \subset \mathcal{U}$ is compact, we can cover it by finitely many $\mathcal{N}_{x_{0,k}}$ as above and find $\varepsilon > 0$ such that $x, x' \in \mathcal{K}$ and $\|x - x'\| < \varepsilon$ is impossible unless x, x' lie in some common $\mathcal{N}_{x_0}$. Consequently there is $c_{\mathcal{K}} > 0$ such that $\|X(t, x') - X(t, x)\| < c_{\mathcal{K}}\|x' - x\|$ whenever $x, x' \in \mathcal{K}$ and $t \in [t_1, t_2]$, because if $\|x - x'\| < \varepsilon$ we can take $c_{\mathcal{K}} \geq \max_k c_{x_{0,k}}$, whereas if $\|x - x'\| \geq \varepsilon$ it is enough to take $c_{\mathcal{K}} > 2M/\varepsilon$ where

M is a bound for $\|X\|$ on $[t_1, t_2] \times \mathcal{K}$. Finally, if $X(t,x)$ happens to vanish identically for x outside some compact $\mathcal{K}' \subset \mathcal{U}$, we can choose $\mathcal{K}$ such that

$$\mathcal{K}' \subset \overset{\circ}{\mathcal{K}} \subset \mathcal{K} \subset \mathcal{U}$$

and construct $c_\mathcal{K}$ as before except that we also pick $\varepsilon > 0$ so small that $\|x - x'\| < \varepsilon$ is impossible for $x \in \mathcal{K}'$ and $x' \notin \mathcal{K}$. Then it holds that $\|X(t,x') - X(t,x)\| < c_\mathcal{K}\|x' - x\|$ for all $x, x' \in \mathcal{U}$ and all $t \in [t_1, t_2]$, that is to say $X(t,x)$ becomes globally Lipschitz with respect to x. These remarks will be used in the proof to come.

Lemma A.3 *Let $t_1 < t_2$ be two real numbers and $X^k : [t_1, t_2] \times \mathcal{U} \to \mathbb{R}^d$ a sequence of time-dependent vector fields, measurable with respect to t, locally Lipschitz continuous with respect to $x \in \mathcal{U}$, and bounded on compact subsets of $[t_1, t_2] \times \mathcal{U}$ independently of k. Let $X : [t_1, t_2] \times \mathcal{U} \to \mathbb{R}^d$ be another time-dependent vector field, measurable with respect to t, locally Lipschitz continuous with respect to $x \in \mathcal{U}$, and assume that, to each compact $\mathcal{K} \subset \mathcal{U}$, there is $E_\mathcal{K} \subset [t_1, t_2]$ of zero measure such that, whenever $t \notin E_\mathcal{K}$, the sequence $X^k(t,x)$ converges to $X(t,x)$ as $k \to \infty$, uniformly with respect to $x \in \mathcal{K}$. Suppose finally that $\gamma : [t_1, t_2] \to \mathcal{U}$ is, for some $(t_0, x_0) \in [t_1, t_2] \times \mathcal{U}$, a solution to the Cauchy problem*

$$\dot\gamma(t) = X(t, \gamma(t)), \qquad \gamma(t_0) = x_0. \tag{A.2}$$

Then, for k large enough, there is a unique solution $\gamma_k : [t_1, t_2] \to \mathcal{U}$ to the Cauchy problem

$$\dot\gamma_k(t) = X^k(t, \gamma_k(t)), \qquad \gamma_k(t_0) = x_0, \tag{A.3}$$

and the sequence (γ_k) converges to γ, uniformly on $[t_1, t_2]$.

Proof. Upon multiplying $X^k(t,x)$ and $X(t,x)$ by a smooth function $\varphi(x)$ which is compactly supported $\mathcal{U} \to \mathbb{R}$ and identically 1 on a neighborhood of $\gamma([t_1, t_2])$, we may assume in view of the discussion preceding the lemma that $X(t,x)$ and $X^k(t,x)$ are defined and bounded $[t_1, t_2] \times \mathbb{R}^d \to \mathbb{R}^d$ independently of k, measurable with respect to t, and (globally) Lipschitz continuous with respect to x.

Then, by classical results [22, Proposition C 3.8., Theorem 54], the solution to (A.3), say γ_k uniquely exists $[t_1, t_2] \to \mathbb{R}^d$ for each k :

$$\gamma_k(t) = x_0 + \int_{t_0}^{t} X^k(s, \gamma_k(s))\, ds, \qquad t \in [t_1, t_2]. \tag{A.4}$$

From the boundedness of X^k, it is clear that γ_k is an equicontinuous and bounded family of functions, hence it is relatively compact in the topology of uniform convergence on $[t_1, t_2]$. All we have to prove then is that every accumulation point of γ_k coincides with γ. Extracting a subsequence if necessary, let us assume that γ_k converges to some $\bar\gamma$, uniformly on $[t_1, t_2]$. Let $\mathcal{K} \subset \mathbb{R}^d$ be a compact set containing $\gamma_k([t_1, t_2])$ for all k; such a set exists by the boundedness of γ_k. If we let $E_\mathcal{K} \subset [t_1, t_2]$ be the set of zero measure granted by the hypothesis, there exists to each $s \in [t_1, t_2] \setminus E_\mathcal{K}$ and each $\varepsilon > 0$ an integer $k_{s,\varepsilon}$ such that

$\|X^k(s,x)-X(s,x)\| < \varepsilon$ as soon as $x \in \mathcal{K}$ and $k > k_{s,\varepsilon}$. In another connection, the Lipschitz character of X with respect to the second argument and the uniform convergence of γ_k to $\bar{\gamma}$ shows that that $\|X(s,\gamma_k(s)) - X(s,\bar{\gamma}(s))\| < \varepsilon$ for k large enough. Altogether, by a 2-ε majorization , we find that

$$\lim_{k\to\infty} \|X^k(s,\gamma_k(s)) - X(s,\bar{\gamma}(s))\| = 0,$$

that is to say the integrand in the right-hand side of (A.4) converges point-wise almost everywhere to $X(s,\bar{\gamma}(s))$. Since X^k is bounded we can apply the dominated convergence theorem and, taking limits on both sides of (A.4) as $k \to \infty$, we find that $\bar{\gamma}$ is a solution to (A.2) whereas the latter is unique. Hence $\bar{\gamma} - \gamma$ as desired. □

B The variational equation

Our goal in this appendix is to give a version of the classical variational equation for ordinary differential equations, in the not-so-classical case where the dependence on time is L^1 but possibly unbounded. Let us first recall the Bellman-Gronwall lemma in a form which is suitable for us.

Lemma B.1 (The Bellman-Gronwall Lemma) *Let w, ϕ, ψ be non-negative real-valued measurable functions on real interval $[0,T]$, such that ψ, ψw and $\psi\phi$ are in $L^1([0,T])$. If it holds that*

$$w(t) \le \phi(t) + \int_0^t \psi(s)w(s)\,ds \quad \text{for } t\in[0,T],$$

then it also holds that

$$w(t) \le \phi(t) + \int_0^t \phi(s)\psi(s)\, \exp\left(\int_s^t \psi(\xi)d\xi\right)\, ds \quad \text{for } t\in[0,T]. \tag{B.1}$$

Proof. By the hypotheses $y(t) = \int_0^t \psi(s)w(s)\,ds$ is an absolutely continuous function of t satisfying

$$\dot{y}(s) - \psi(s)y(s) \le \phi(s)\psi(s) \qquad \text{a. e. } s\in[0,T],$$

therefore $z(t) = y(t)\exp(-\int_0^t \psi(s)\,ds)$ is also absolutely continuous and satisfies

$$\dot{z}(s) \le \phi(s)\psi(s)\exp\left(-\int_0^s \psi(\tau)\,\mathrm{d}\tau\right) \qquad \text{a. e. } s\in[0,T]. \tag{B.2}$$

Integrating (B.2) from 0 to t and multiplying by $\exp(\int_0^t \psi(s)\,ds)$ yields

$$y(t) \le \int_0^t \phi(s)\psi(s)\, \exp\left(\int_s^t \psi(\xi)d\xi\right)\, ds \quad \text{for } t\in[0,T],$$

from which (B.1) follows since $w(t) \leq \phi(t) + y(t)$ by hypothesis, compare for instance [7, sec. 10.5.1.3]. □

Let us now consider a differential equation of the form

$$\dot{x} = X(x,t) \tag{B.3}$$

where the time-dependent vector field $X : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ satisfies the following properties:

(i) for fixed $t \in \mathbb{R}$, the map $x \to X(x,t)$ is continuously differentiable $\mathbb{R}^n \to \mathbb{R}^n$;

(ii) for fixed $x \in \mathbb{R}^n$, the map $t \to X(x,t)$ is measurable $\mathbb{R} \to \mathbb{R}$;

(iii) for some $x_1 \in \mathbb{R}^n$ there is a measurable and locally integrable function $\alpha_{x_1} : \mathbb{R} \to \mathbb{R}^+$ such that
$$\|X(x_1,t)\| \leq \alpha_{x_1}(t), \qquad \text{for all } t \in \mathbb{R};$$

(iv) there is a measurable and locally integrable function $\psi : \mathbb{R} \to \mathbb{R}^+$ satisfying
$$\left\| \frac{\partial X}{\partial x}(x,t) \right\|_{\mathrm{O}} \leq \psi(t), \qquad \text{for all } (x,t) \in \mathbb{R}^n \times \mathbb{R},$$
where $\| \ \|_{\mathrm{O}}$ denotes the familiar operator norm on $n \times n$ real matrices.

The choice of the operator norm in **(iv)** is only for definiteness since all norms are equivalent on $\mathbb{R}^{n\times n}$. Note also that, using **(iv)** and the mean-value theorem, property **(iii)** immediately strengthens to:

(iii)' to each $x \in \mathbb{R}^n$ there is a measurable and locally integrable function $\alpha_x : \mathbb{R} \to \mathbb{R}^+$ such that
$$\|X(x,t)\| \leq \alpha_x(t), \qquad \text{for all } t \in \mathbb{R}.$$

By **(i)**, **(ii)**, **(iii)'**, and **(iv)**, the solution to (B.3) with arbitrary initial condition $x(0) = x_0 \in \mathbb{R}^n$ uniquely exists for all $t \in \mathbb{R}$, *cf.* [22, Theorem 54, Proposition C.3.4, Proposition C.3.8], in the sense that there is a unique locally absolutely continuous function $x : \mathbb{R} \to \mathbb{R}^n$ satisfying (B.3) for almost every t and such that $x(0) = x_0$. We shall denote by $\widehat{x}(\tau, x_0)$ the value of this solution at time $t = \tau$, in other words we let $(t, x_0) \mapsto \widehat{x}(\tau, x_0)$ designate the flow of (B.3). By definition, the *variational equation* of (B.3) along the trajectory $t \mapsto \widehat{x}(t, x_0)$ is the linear differential equation:

$$\dot{R} = \frac{\partial X}{\partial x}(\widehat{x}(t,x_0),t)\,R \tag{B.4}$$

in the unknown matrix-valued function $R : \mathbb{R} \to \mathbb{R}^{n\times n}$. In view of **(iv)**, appealing again to [22, Theorem 54, Proposition C.3.4, Proposition C.3.8], we see that the solution to (B.4) uniquely exists for all t once some arbitrary initial condition $R(0) = R_0 \in \mathbb{R}^{n\times n}$ is prescribed. Accordingly, we let $\widehat{R}(t, R_0, x_0)$ denote the value at time t of that solution.

Proposition B.2 *If $X : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ satisfies properties* **(i)**-**(iv)** *above, and if $\widehat{x}$, $\widehat{R}$ are the respective flows of* (B.3), (B.4) *defined previously, then $\widehat{x}(t,x)$ is continuously differentiable with respect to x for fixed t and*

$$\frac{\partial \widehat{x}}{\partial x}(t,x) \;=\; \widehat{R}(t, I_n, x), \tag{B.5}$$

where I_n is the identity matrix of size n.

Proof. Upon changing $X(x,t)$ into $-X(x,-t)$ if necessary, we may assume throughout the proof that $t \geq 0$. We first show that $x \mapsto \widehat{x}(t,x)$ is continuous for fixed t. Indeed, setting for $x, h \in \mathbb{R}^n$

$$\delta(t,x,h) \stackrel{\Delta}{=} \widehat{x}(t,x+h) - \widehat{x}(t,x),$$

we get by definition of $\widehat{x}$ that $\delta(t,x,h)$ is locally absolutely continuous with respect to t for fixed x, h, with derivative given almost everywhere by

$$\begin{aligned} \dot{\delta}(t,x,h) \;&=\; X(\widehat{x}(t,x+h),t) - X(\widehat{x}(t,x),t) \\ &=\; \left(\int_0^1 \frac{\partial X}{\partial x}\Big(\tau\widehat{x}(t,x+h) + (1-\tau)\widehat{x}(t,x), t\Big)\,\mathrm{d}\tau\right)\delta(t,x,h), \end{aligned} \tag{B.6}$$

where we have used point **(i)** of our hypotheses. If we put for simplicity :

$$T(x,h,s) \stackrel{\Delta}{=} \int_0^1 \frac{\partial X}{\partial x}\Big(\tau\widehat{x}(s,x+h) + (1-\tau)\widehat{x}(s,x)\,,\; s\Big)\,\mathrm{d}\tau \tag{B.7}$$

and if we notice by point **(iv)** of the hypotheses that

$$\|T(x,h,s)\|_{\mathrm{O}} \leq \psi(s), \tag{B.8}$$

we deduce from (B.6) and (B.8), since $\delta(0,x,h) = h$, that

$$\|\delta(t,x,h)\| \leq \|h\| + \int_0^t \psi(s)\,\|\delta(s,x,h)\|\,ds.$$

As ψ is locally L^1 while $s \mapsto \delta(s,x,h)$ is *a fortiori* continuous hence bounded on $[0,t]$, Lemma B.1 implies that

$$\|\delta(t,x,h)\| \leq \|h\|\left(1 + \exp\left(\int_0^t \psi(\xi)\,d\xi\right)\int_0^t \psi(s)\,ds\right). \tag{B.9}$$

Since the right-hand side of (B.9) can be made arbitrarily small with $\|h\|$, we get the announced continuity of $x \mapsto \widehat{x}(t,x)$.

Next, we put for $x, h \in \mathbb{R}^n$

$$y(t,x,h) \;\stackrel{\Delta}{=}\; \widehat{x}(t,x+h) - \widehat{x}(t,x) - \widehat{R}(t, I_n, x)\,h, \tag{B.10}$$

and we need to show that $\|y(t,x,h)\|$ is little $o(\|h\|)$ for fixed t,x. Clearly $t \mapsto y(t,x,h)$ is locally absolutely continuous with $y(0,x,h)=0$. Computing its derivative using (B.10), (B.3), and (B.4), we get

$$\begin{aligned} y(t,x,h) \quad = \quad & \int_0^t \Big(X(\widehat{x}(s,x+h),s) - X(\widehat{x}(s,x),s) \\ & \qquad - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s)\,\widehat{R}(s,I_n,x)\,h \Big)\,ds. \end{aligned} \tag{B.11}$$

In view of (B.10), making use of the second equality in (B.6), we may rewrite (B.11) in the form :

$$\begin{aligned} y(t,x,h) \quad = \quad & \int_0^t T(x,h,s)\,y(s,x,h)\,ds \\ & + \int_0^t \left(T(x,h,s) - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s) \right) \widehat{R}(s,I_n,x)\,h\,ds \end{aligned} \tag{B.12}$$

where $T(x,h,s)$ was defined in (B.7). If we further define

$$\Phi(t,x,h) \triangleq \left\| \int_0^t \left(T(x,h,s) - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s) \right) \widehat{R}(s,I_n,x)\,ds \right\|_{\mathrm{O}} \tag{B.13}$$

we obtain from (B.12) and (B.8) the inequality :

$$\|y(t,x,h)\| \le \Phi(t,x,h)\,\|h\| + \int_0^t \psi(s)\,\|y(s,x,h)\|\,ds.$$

Observe by (B.8) and point **(iv)** of the hypotheses that

$$\left\| T(x,h,s) - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s) \right\|_{\mathrm{O}} \le 2\psi(s), \tag{B.14}$$

so that $t \mapsto \Phi(t,x,h)$ is locally bounded for fixed x, uniformly with respect to $h \in I\!\!R^n$, because ψ is locally L^1 and because $s \mapsto \|\widehat{R}(s,I_n,x)\|_{\mathrm{O}}$ is continuous hence locally bounded. Since $t \mapsto y(t,x,h)$ is also continuous hence locally bounded, Lemma B.1 yields :

$$\|y(t,x,h)\| \le \Phi(t,x,h)\,\|h\| + \|h\| \exp\left(\int_0^t \psi(\xi)\,d\xi \right) \int_0^t \psi(s)\Phi(s,x,h)\,ds.$$

From this, appealing to the dominated convergence theorem, we shall deduce that $\|y(t,x,h)\|$ is little $o(\|h\|)$ for fixed t,x if only we can show that $s \mapsto \Phi(s,x,h)$ goes boundedly point-wise to zero with $\|h\|$ on $[0,t]$. In fact, we just pointed out that it is bounded there, independently of h. To see that it converges point-wise to zero when $\|h\| \to 0$, we return to the definition (B.13) of Φ and, taking into account (B.14) where ψ is locally L^1 and the already used boundedness of $s \mapsto \|\widehat{R}(s, I_n, x)\|_{\mathrm{O}}$ on $[0,t]$ for fixed x, we observe that it is enough by dominated convergence to establish the point-wise limit :

$$\lim_{\|h\|\to 0} \left\| T(x,h,s) - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s) \right\| = 0, \qquad x \in I\!\!R^n,\ s \in [0,t].$$

The latter in turn follows from another application of the dominated convergence theorem to the right-hand side of (B.7), considering points **(i)** and **(iv)** of the hypotheses together with the continuity of $x \mapsto \widehat{x}(t,x)$ proved earlier.

To complete the proof, it remains for us to show that $x \mapsto \widehat{R}(t, I_n, x)$ is continuous for fixed t. In other words, if we put for $x, h \in I\!\!R^n$:

$$\Delta(t,x,h) \triangleq \widehat{R}(t, I_n, x+h) - \widehat{R}(t, I_n, x),$$

we need to show that $\|\Delta(t,x,h)\|_{\mathrm{O}}$ is little $o(\|h\|)$ as $\|h\| \to 0$ for fixed t and x. To this effect, using (B.4), we write

$$\begin{aligned}
\Delta(t,x,h) &= \int_0^t \left(\frac{\partial X}{\partial x}(\widehat{x}(s,x+h),s)\, \widehat{R}(s, I_n, x+h) \right.\\
&\qquad\qquad \left. - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s)\, \widehat{R}(s, I_n, x) \right) ds \\
&= \int_0^t \left(\frac{\partial X}{\partial x}(\widehat{x}(s,x+h),s)\, \Delta(s,x,h) \right. \qquad\qquad (B.15)\\
&\qquad \left. + \left(\frac{\partial X}{\partial x}(\widehat{x}(s,x+h),s) - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s) \right) \widehat{R}(s, I_n, x) \right) ds.
\end{aligned}$$

Setting

$$\Theta(t,x,h) \triangleq \left\| \int_0^t \left(\frac{\partial X}{\partial x}(\widehat{x}(s,x+h),s) - \frac{\partial X}{\partial x}(\widehat{x}(s,x),s) \right) \widehat{R}(s, I_n, x)\, ds \right\|_{\mathrm{O}},$$

we obtain from (B.15) and point **(iv)** of the hypotheses that

$$\|\Delta(t,x,h)\|_{\mathrm{O}} \leq \int_0^t \psi(s) \|\Delta(s,x,h)\|_{\mathrm{O}}\, ds + \Theta(t,x,h). \tag{B.16}$$

Since $t \to \Theta(t,x,h)$ is locally bounded for fixed x independently of $h \in I\!\!R^n$, as follows from point **(iv)** again and the fact that $s \mapsto \widehat{R}(s, I_n, x)$ is continuous hence bounded on $[0,t]$, Lemma B.1 now yields :

$$\|\Delta(t,x,h)\|_{\mathrm{O}} \leq \Theta(t,x,h) + \exp\left(\int_0^t \psi(\xi)\, d\xi \right) \int_0^t \psi(s) \Theta(s,x,h)\, ds.$$

From this, appealing to the dominated convergence theorem, we shall deduce that $\|\Delta(t,x,h)\|_{\mathrm{O}}$ is little $o(\|h\|)$ for fixed t,x if only we can show that $s \mapsto \Theta(s,x,h)$ goes boundedly pointwise to zero with $\|h\|$ on $[0,t]$. But we already proved its boundedness, and the desired limit :

$$\lim_{\|h\|\to 0} \Theta(t,x,h) = 0, \qquad x \in I\!R^n,\ t \in I\!R,$$

follows from yet another application of the dominated convergence theorem in the equation defining Θ, granted points **(i)** and **(iv)** of the hypotheses together with the continuity of $x \mapsto \widehat{x}(t,x)$ already established. $\square$

C Continuity of the flow with L^p controls

In this appendix, we deal with a differential equation of the form

$$\dot{x} = F(x, \Upsilon(t)) \tag{C.1}$$

where $x \in I\!R^n$ while $\Upsilon \in L^p = L^p(I\!R, I\!R^m)$, the familiar Lebesgue space of (equivalence classes of) functions $I\!R \to I\!R^m$ whose p-th power is integrable in case $p < \infty$ and whose norm is essentially bounded if $p = \infty$; we endow L^p with the usual norm, namely $\|\Upsilon\|_p = (\int_{I\!R} \|\Upsilon\|^p dt)^{1/p}$ if $p < \infty$ and $\|\Upsilon\|_\infty = \text{ess.sup.}_{I\!R}\|\Upsilon\|$, where $\|.\|$ denotes the Euclidean norm. Of course, a solution to the differential equation is understood here in the sense that $x(t)$ is absolutely continuous, and that its derivative is a locally summable function whose value is given by the right-hand side of (C.1) for almost every t. Classically, even if $F : I\!R^n \times I\!R^m \to I\!R^n$ is very smooth, the existence of solutions to (C.1) when $1 \le p < \infty$ requires some restrictions on the growth of F at infinity. Even then however, the continuity of that solution with respect to $\Upsilon \in L^p$ is difficult to ferret out in the literature. We propose below a set of conditions that ensures such a continuity property, this result being used in the proof of Theorem 7.10. For definiteness, we agree in the statement that $\|.\|$ refers to the operator norm when applied to a matrix.

Proposition C.1 *Let $F(x,u)$ be continuous $I\!R^n \times I\!R^m \to I\!R^n$, and the partial derivative $\partial F/\partial x$ exist continuously $I\!R^n \times I\!R^m \to I\!R^{n\times n}$. Let $p \in [1,\infty]$ and assume if $p < \infty$ that, to each compact $K \subset I\!R^n$, there are constants $c_1(K)$, $c_2(K)$, such that :*

$$\|F(x,u)\| + \|\frac{\partial F}{\partial x}(x,u)\| \le c_1(K) + c_2(K)\,\|u\|^p, \quad (x,u) \in K \times I\!R^m. \tag{C.2}$$

Then, for any $\Upsilon \in L^p(I\!R, I\!R^m)$, the solution $t \mapsto x(t, x_0, \Upsilon)$ to (C.1) with initial condition $x(0) = x_0$ uniquely exists on some maximal time interval $\mathcal{I}_{x_0,\Upsilon}$ containing 0. Moreover, if $\mathcal{K}$ is a compact subinterval of $\mathcal{I}_{x_0,\Upsilon}$, there is a neighborhood $\mathcal{V}$ of (x_0,Υ) in $I\!R^n \times L^p(I\!R, I\!R^m)$ such that $\mathcal{K} \subset \mathcal{I}_{x_0',\Upsilon'}$ whenever $(x_0', \Upsilon') \in \mathcal{V}$; within this neighborhood, it further holds that

$$\lim_{(x_0',\Upsilon')\to(x_0,\Upsilon)} x(t, x_0', \Upsilon') = x(t, x_0, \Upsilon), \tag{C.3}$$

uniformly with respect to $t \in \mathcal{K}$.

Proof. If $\Upsilon \in L^p$, and provided (C.2) holds in case $p < \infty$, it follows immediately from classical existence and uniqueness results (see *e.g.* [22, Theorem 54, Proposition C.3.4])[5] that $x(t, x_0, \Upsilon)$ is uniquely defined on some maximal time interval containing 0, say $\mathcal{I}_{x_0,\Upsilon}$. Next, let us replace $F(x,u)$ by $F_1(x,u) = \varphi(x)F(x,u)$, where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is smooth with compact support and assumes the value 1 on a neighborhood of the compact set $x(\mathcal{K}, x_0, \Upsilon)$. Note that F_1 again satisfies an estimate of the form (C.2) if F does, and that it vanishes for x outside the support of φ. Therefore, if F gets replaced by F_1, (C.2) will hold when $p < \infty$ for some constants c_1, c_2 that are in fact independent of K, whereas if $p = \infty$ $\partial F_1/\partial x\,(x, \Upsilon(t))$ is bounded by a constant a.e. in t for fixed $\Upsilon \in L^\infty$. This is to the effect that, if we deal with F_1 instead of F, the solution to (C.1) exists for all $t \in \mathbb{R}$ [22, Proposition C.3.8]. This entails that if we prove the proposition for F_1, then we get it for F as well, because the property for system (C.1) that $\mathcal{K} \subset \mathcal{I}_{x_0',\Upsilon'}$ whenever (x_0', Υ') is sufficiently close to (x_0, Υ) in $\mathbb{R}^n \times L^p$ will be a mechanical consequence of property (C.3) for the system $\dot{x} = F_1(x, \Upsilon(t))$, granted that $F(x,u)$ and $F_1(x,u)$ coincide for x in a neighborhood of $x(\mathcal{K}, x_0, \Upsilon)$. To recap, we are left to prove (C.3) under the stronger assumption that $F(x,u)$ hence also $\partial F/\partial x$ vanishes for x outside some compact set, in which case $c_1(K)$ and $c_2(K)$ in (C.2) are taken to be absolute constants c_1 and c_2, while $\mathcal{I}_{x_0,\Upsilon} = \mathbb{R}$ for all $(x_0, \Upsilon) \in \mathbb{R}^n \times L^p$.

Pick $(x_0', \Upsilon') \in \mathbb{R}^n \times L^p$ and set for simplicity $x(t) = x(t, x_0, \Upsilon)$ and $x'(t) = x(t, x_0', \Upsilon')$. From the definitions, we get that

$$\begin{aligned} x(t) - x'(t) &= x_0 - x_0' + \int_0^t \left(F\left(x(\tau), \Upsilon(\tau)\right) - F\left(x(\tau), \Upsilon'(\tau)\right)\right) \mathrm{d}\tau \\ &\quad + \int_0^t \left(F\left(x(\tau), \Upsilon'(\tau)\right) - F\left(x'(\tau), \Upsilon'(\tau)\right)\right) \mathrm{d}\tau. \end{aligned}$$

If $p = \infty$, we obtain at once from the mean-value theorem :

$$\begin{aligned} \|x(t) - x'(t)\| &\leq \|x_0 - x_0'\| + \int_0^t \|F(x(\tau), \Upsilon(\tau)) - F((x(\tau), \Upsilon'(\tau))\|\, \mathrm{d}\tau \\ &\quad + \sup_{\substack{(x,u) \in \mathbb{R}^n \times \mathbb{R}^m \\ \|u\| \leq \|\Upsilon'\|_\infty}} \left\| \frac{\partial F}{\partial x}(x,u) \right\| \int_0^t \|x(\tau) - x'(\tau)\|\, \mathrm{d}\tau, \end{aligned} \tag{C.4}$$

and if $1 \leq p < \infty$ we additionally take (C.2) into account to get :

$$\begin{aligned} \|x(t) - x'(t)\| &\leq \|x_0 - x_0'\| + \int_0^t \|F(x(\tau), \Upsilon(\tau)) - F((x(\tau), \Upsilon'(\tau))\|\, \mathrm{d}\tau \\ &\quad + \int_0^t \left(c_1 + c_2 \|\Upsilon'(\tau)\|^p\right) \|x(\tau) - x'(\tau)\|\, \mathrm{d}\tau. \end{aligned} \tag{C.5}$$

[5] Strictly speaking, to apply Theorem 54 of that reference, we need to choose a specific representative of Υ which is defined *everywhere*; this causes no difficulty because the solution of course does not depend on this representative.

To establish (C.3), we may of course assume that $\|\Upsilon'\|_p$ remains bounded and therefore, by the Bellman-Gronwall lemma as applied to (C.4) or (C.5) according whether $p=\infty$ or $p<\infty$ (see Lemma B.1), we shall be done if only we can show that

$$\phi_{\Upsilon'}(t)=\int_0^t \|F(x(\tau),\Upsilon(\tau))-F((x(\tau),\Upsilon'(\tau))\|\,\mathrm{d}\tau, \tag{C.6}$$

can be made small with $\|\Upsilon'-\Upsilon\|_p$ for fixed $t\in I\!R$ (compare [22, Theorem 55]). This is obvious if $p=\infty$ by the uniform continuity of F relatively to the compact set $x([0,t])\times\overline{B}(0,\|\Upsilon\|_\infty)$, thus we assume in the remaining of the proof that $p<\infty$. Choose Υ' such that $\|\Upsilon'-\Upsilon\|_p<\varepsilon$. Since both $\|\Upsilon\|^p$ and $F(x(\tau),\Upsilon(\tau))$ are summable using (C.2), there is by absolute continuity an $\eta>0$ such that

$$\max\left\{\int_E \|\Upsilon(\tau)\|^p\,\mathrm{d}\tau\ ,\ \int_E \|F(x(\tau),\Upsilon(\tau))\|\,\mathrm{d}\tau\right\}<\varepsilon \quad \text{whenever } |E|<\eta,$$

where $|E|$ denotes the Lebesgue measure of a measurable set $E\subset I\!R$ [21, Theorem 6.11]. Then, again from (C.2), we have that

$$\begin{aligned}\int_E \|F(x(\tau),\Upsilon'(\tau))\|\,\mathrm{d}\tau &\le c_1|E|+c_2\int_E \|\Upsilon'(\tau)\|^p\,\mathrm{d}\tau\\ &\le c_1|E|+c_2'\int_E \left(\|\Upsilon'(\tau)-\Upsilon(\tau)\|^p+\|\Upsilon(\tau)\|^p\right)\mathrm{d}\tau\end{aligned}$$

for some constant c_2' ($c_2 2^{p/q}$ will do if $1/p+1/q=1$).

Using the triangle inequality and collecting terms, we find that

$$\int_E \|F(x(\tau),\Upsilon(\tau))-F((x(\tau),\Upsilon'(\tau))\|\,\mathrm{d}\tau\le c_1\eta+\varepsilon(1+c'2+c_2'\varepsilon^{p-1}),$$

and if we further impose, without loss of generality, that $\eta\le\varepsilon<1$ while putting $c_3=1+c_1+2c_2'$, we obtain :

$$\int_E \|F(x(\tau),\Upsilon(\tau))-F((x(\tau),\Upsilon'(\tau))\|\,\mathrm{d}\tau\le c_3\,\varepsilon \quad \text{if } |E|<\eta. \tag{C.7}$$

Now, pick $M>0$ so large that

$$E_M=\{\tau\in I\!R;\ \|\Upsilon(\tau)\|>M\} \tag{C.8}$$

has Lebesgue measure less than $<\eta$. By uniform continuity of F relatively to $x([0,t])\times\overline{B}(0,M)$, there is $\alpha>0$ such that

$$\|F(x(\tau),u')-F(x(\tau),u)\|<\varepsilon \ \text{ for } \tau\in[0,t],\ \|u\|\le M,\ \|u'-u\|<\alpha.$$

Let us further define

$$E_{\alpha,\Upsilon'}=\{\tau\in I\!R;\ \|\Upsilon'(\tau)-\Upsilon(\tau)\|\ge\alpha\}. \tag{C.9}$$

By (C.8), (C.9), and the definition of α, we get that

$$\|F(x(\tau), \Upsilon(\tau)) - F(x(\tau), \Upsilon'(\tau))\| < \varepsilon \quad \text{for } \tau \in [0,t] \setminus (E_M \cup E_{\alpha,\Upsilon'}). \tag{C.10}$$

Finally, since $|E_{\alpha,\Upsilon'}| \leq \|\Upsilon - \Upsilon'\|_p/\alpha$, we can make it less than η by requiring that $\|\Upsilon - \Upsilon'\|_p < \eta\alpha$. Altogether, starting from $0 < \varepsilon < 1$, we have found $\eta > 0$ and $\alpha > 0$ such that, if

$$\|\Upsilon - \Upsilon'\|_p < \max\{\eta\alpha, \varepsilon\},$$

then both E_M defined by (C.8) and $E_{\alpha,\Upsilon'}$ defined by (C.9) have Lebesgue measure less than η while (C.7) and (C.10) hold. When these conditions are satisfied, we get upon decomposing

$$\int_{[0,t]} = \int_{E_M} + \int_{E_{\alpha,\Upsilon'}} + \int_{[0,t]\setminus(E_M \cup E_{\alpha,\Upsilon'})}$$

that $\phi_{\Upsilon'}(t)$ defined in (C.6) is less than $\varepsilon(|t| + 2c_3)$ which is arbitrarily small, as announced. □

D Orbits of families of vector fields

In the proof of lemma F.1 we will need results from [24] on orbits[6] of families of smooth vector fields, that were recently exposed in the textbook [13, chapter II]. We recall them below, in a slightly expanded form.

Let $\mathcal{F}$ be a family of smooth vector fields defined on an open subset U of $\mathbb{R}^d$. For any positive integer N and vector fields $X^1, \ldots, X^N$ belonging to $\mathcal{F}$, given $m \in U$, consider the map F given by

$$(t_1, \ldots, t_N) \quad \mapsto \quad X^1_{t_1}(X^2_{t_2}(\cdots(X^N_{t_N}(m))\cdots)) \tag{D.1}$$

where the standard notation $X_t(x)$ indicates the flow of X from x at time t; of course, F depends on the choice of the vector fields X^j and of the point m. This map is defined on some open connected neighborhood of the origin, hereafter denoted by $\mathrm{dom}(F)$, and takes values in U. In fact, $(t_1, \ldots, t_N) \in \mathrm{dom}(F)$ if, and only if, for every $j \in \{1, \ldots, N\}$, the solution $x(\tau)$ to $\dot{x} = X^j(x)$, with initial condition $x(0) = X^{j-1}_{t_{j-1}}(\cdots(X^1_{t_1}(m))\cdots)$, exists in U for all $\tau \in [0, t_j]$ (or $[t_j, 0]$ if $t_j < 0$).

The *orbit* of the family $\mathcal{F}$ through a point $m \in U$ is the set of all points that lie in the image of F for at least one choice of the vector fields $X^1, \ldots, X^N$. In words, the orbit of the family $\mathcal{F}$ through m is the set of points that may be linked to m in U upon concatenating finitely many integral curves of vector fields in the family. We shall denote by $\mathcal{O}_{\mathcal{F},p}$ the orbit of $\mathcal{F}$ through m.

[6] One of the motivations in [24] was to generalize the notion of integral manifolds to vector fields that are smooth but not real analytic. Note that the orbits of a family of *real analytic* vector fields actually coincide with the maximal integral manifolds of the closure of this family under Lie brackets [24, 15, 17]. However, even if we assume the control system (3.1) to be real analytic, integral manifolds are of no help to us because topological conjugacy does not preserve tangency nor Lie brackets. Using orbits of families of vector fields instead is much more efficient, because topological conjugacy does preserve integral curves.

Note that the definition depends on U in a slightly subtle manner : if $\mathcal{F}$ defines by restriction a family of vector fields $\mathcal{F}|_V$ on a smaller open set $V \subset U$ and if $m \in V$, then

$$V \cap \mathcal{O}_{\mathcal{F},m} \quad \supset \quad \mathcal{O}_{\mathcal{F}|_V,m}, \tag{D.2}$$

but the inclusion is generally strict because of the requirement that the integral curves used to construct $\mathcal{O}_{\mathcal{F}|_V,m}$ should lie entirely in V.

We turn to topological considerations. The topology of U is the usual Euclidean topology. The topology of $\mathcal{O}_{\mathcal{F},m}$ *as an orbit* is the finest that makes all the maps F, arising from (D.1), continuous on their respective domains of definition, the latter being endowed with the Euclidean topology. The classical smoothness of the flow implies that each F is continuous $\mathrm{dom}(F) \to I\!R^d$, hence the topology of $\mathcal{O}_{\mathcal{F},m}$ as an orbit is finer than the Euclidean topology induced by the ambient space U. *It can be strictly finer*, and this is why we speak of the *orbit topology*, as opposed to the *induced topology*.

Starting from $\mathcal{F}$, one defines a larger family of vector fields $P_\mathcal{F}$, consisting of all the push-forwards[7] of vector fields in $\mathcal{F}$ through all local diffeomorphisms of the form $X^1_{t_1} \circ X^2_{t_2} \circ \cdots \circ X^N_{t_N}$ where $X^1, \ldots, X^N$ belong to $\mathcal{F}$. That is to say, vector fields in $P_\mathcal{F}$ are of the form

$$\left(X^1_{t_1} \circ \cdots \circ X^N_{t_N}\right)_\star X^0 \tag{D.3}$$

where $X^0, X^1, \ldots, X^N$ belong to $\mathcal{F}$.

Remark D.1 *Note that a member of $P_\mathcal{F}$ is defined on an open set which is generally a strict subset of U, whereas members of $\mathcal{F}$ are defined over the whole of U, and it is understood that a curve $\gamma : I \to U$, where I is a real interval, will be called an integral curve of $Y \in P_\mathcal{F}$ only when $\gamma(I)$ is included in the domain of definition of Y.*

For $x \in U$, we denote by $P_\mathcal{F}(x)$ the subspace of $I\!R^d$ spanned by all the vectors $Y(x)$, where $Y \in P_\mathcal{F}(x)$ is defined in a neighborhood of x.

Theorem D.2 below, which is the central result in this appendix, describes the topological nature of the orbits. To interpret the statement correctly, it is necessary to recall (see for instance [23]) that an *immersed* sub-manifold of a manifold is a subset of the latter which is a manifold in its own right, and is such that the inclusion map is an immersion. This allows one to naturally identify the tangent space to an immersed sub-manifold at a given point with a linear subspace of the tangent space to the ambient manifold at the same point. The topology of an immersed sub-manifold is in general finer than the one induced by the ambient manifold; when these two topologies coincide, the sub-manifold is called *embedded*.

Theorem D.2 (Orbit Theorem, Sussmann [24]) *Let $\mathcal{F}$ be a family of smooth vector fields defined on an open set $U \subset I\!R^d$, and m be a point in U. If $\mathcal{O}_{\mathcal{F},m}$ denotes the orbit of $\mathcal{F}$ through m, then:*

[7]Recall that the push-forward of a vector field $X : V \to I\!R^d$ through a diffeomorphism $\varphi : V \to \varphi(V)$ is the vector field $\varphi_\star X$ on $\varphi(V)$ whose flow at each time is the conjugate of the flow of X under the diffeomorphism φ; it can be defined as $\varphi_\star X(\varphi(x)) = D\varphi(x)X(x)$, where $D\varphi(x)$ is the derivative of φ at $x \in V$.

(i) *Endowed with the orbit topology, $\mathcal{O}_{\mathcal{F},m}$ has a unique differential structure that makes it a smooth connected immersed sub-manifold of U, for which the maps* (D.1) *are smooth.*

(ii) *The tangent space to $\mathcal{O}_{\mathcal{F},m}$ at $x \in \mathcal{O}_{\mathcal{F},m}$ is $P_{\mathcal{F}}(x)$.*

(iii) *There exists an open neighborhood W of m in U, and smooth local coordinates $\xi : W \to (-\eta, \eta)^d \subset \mathbb{R}^d$, with $\xi(m) = 0$, such that*

(a) *in these coordinates, $W \cap \mathcal{O}_{\mathcal{F},m}$ is a product :*

$$W \cap \mathcal{O}_{\mathcal{F},m} = (-\eta, \eta)^q \times T \tag{D.4}$$

where $\eta > 0$, q is the dimension of $\mathcal{O}_{\mathcal{F},m}$, and T is some subset of $(-\eta, \eta)^{d-q}$ containing the origin. The orbit topology of $\mathcal{O}_{\mathcal{F},m}$ induces on $W \cap \mathcal{O}_{\mathcal{F},m}$ the product topology where $(-\eta, \eta)^q$ is endowed with the usual Euclidean topology and T with the discrete topology.

(b) *if $\gamma : [t_1, t_2] \to W \cap \mathcal{O}_{\mathcal{F},m}$ is an integral curve of a vector field $Y \in P_{\mathcal{F}}$ (see remark D.1), then $t \mapsto \xi_i(\gamma(t))$, $q + 1 \le i \le d$, are constant mappings,*

(c) *the tangent space to $\mathcal{O}_{\mathcal{F},m}$ at each point $p \in W \cap \mathcal{O}_{\mathcal{F},m}$ is spanned by the vector fields $\partial/\partial\xi_1, \ldots, \partial/\partial\xi_q$,*

(d) *at any point $p \in W$, the vector fields $\partial/\partial\xi_1, \ldots, \partial/\partial\xi_q$ belong to the tangent space to the orbit of $\mathcal{F}$ through p.*

Remark D.3 *Another description of the product topology in point $(iii) - (a)$ is as follows. The connected components of $W \cap \mathcal{O}_{\mathcal{F},m}$ are the sets*

$$S_{W,a} = (-\eta, \eta)^q \times \{a\} \tag{D.5}$$

for $a \in T$, and the topology on each of these connected components is the topology induced by the ambient Euclidean topology. In particular each $S_{W,a}$ is an embedded *sub-manifold of U.*

Proof of Theorem D.2. Assertion (i) is the standard form of the orbit theorem (cf *e.g.* [13, Chapter 2, Theorem 1]), while assertion (ii) is a rephrasing of [24, Theorem 4.1, point (b)]. Assertion (iii) apparently cannot be referenced exactly in this form, but we shall deduce it from the previous ones as follows.

By point (ii), the tangent space to $\mathcal{O}_{\mathcal{F},m}$ at $m \in S$ is the linear span over $\mathbb{R}$ of $Y^1(m), \ldots, Y^q(m)$, where $Y^1, \ldots, Y^q$ are q vector fields belonging to $P_{\mathcal{F}}$, defined on some neighborhood of m, and such that $Y^1(m), \ldots, Y^q(m)$ are linearly independent (recall that q is the dimension of $\mathcal{O}_{\mathcal{F},m}$). Let us write

$$Y^j = \left(X^{j,1}_{t_{j,1}} \circ \cdots \circ X^{j,N_j}_{t_{j,N_j}} \right)_\star X^{j,0}, \quad 1 \le j \le q,$$

where $X^{j,k} \in \mathcal{F}$ for $0 \le k \le N_j$, and where the $t_{j,k}$'s are real numbers for which the concatenated flow exists, locally around m (compare (D.3)).

Since $Y^1(m),\ldots,Y^q(m)$ are linearly independent, one may complement them into a basis of $\mathbb{R}^d$ by adjunction of $d-q$ independent vectors that may, without loss of generality, be regarded as values at m of $d-q$ smooth vector fields in U, say $Y^{q+1},\ldots,Y^d$. Then, the smooth map

$$L(\xi_1,\ldots,\xi_d) \;=\; \left(Y^1_{\xi_1}\circ\cdots\circ Y^q_{\xi_q}\circ Y^{q+1}_{\xi_{q+1}}\circ\cdots\circ Y^d_{\xi_d}\right)(m) \tag{D.6}$$

defines a diffeomorphism from some poly-interval $\mathcal{I}_\eta=\{(\xi_1,\ldots,\xi_d)\,,\;|\xi_i|<\eta\}$ onto an open neighborhood W of m in U, simply because the derivative of L is invertible at the origin as $Y^1(m),\ldots,Y^d(m)$ are linearly independent by construction. Let $\xi:W\to\mathcal{I}_\eta$ denote its inverse.

By the characteristic property of push-forwards, we locally have, for $1\le j\le q$, that

$$Y^j_{\xi_j}=X^{j,1}_{t_{j,1}}\circ\cdots\circ X^{j,N_j}_{t_{j,N_j}}\circ X^{j,0}_{\xi_j}\circ X^{j,N_j}_{-t_{j,N_j}}\circ\cdots\circ X^{j,1}_{-t_{j,1}}\,. \tag{D.7}$$

This implies that, in (D.6), the images under L of those d-tuples sharing a common value of $\xi_{q+1},\ldots,\xi_d$ all lie in the same orbit $\mathcal{O}_{\mathcal{F},L(0,\ldots,0,\xi_{q+1},\ldots,\xi_d)}$. In particular, the map

$$\tau_1,\ldots,\tau_q \;\mapsto\; \left(Y^1_{\tau_1+\xi_1}\circ\cdots\circ Y^q_{\tau_q+\xi_q}\circ Y^{q+1}_{\xi_{q+1}}\circ\cdots\circ Y^d_{\xi_d}\right)(m)$$

is defined $\Pi_{j=1}^q(-\eta-\xi_j,\eta-\xi_j)\to W\cap\mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}$, and this map is smooth from the Euclidean to the orbit topology by (D.7) and point (i). If we compose it with the immersive injection $J_W:W\cap\mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}\to W$ (keeping in mind that $W\cap\mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}$ is open in $\mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}$ since the orbit topology is finer than the Euclidean one), and if we subsequently apply ξ, we get the affine map

$$\tau_1,\ldots,\tau_q \;\mapsto\; (\tau_1+\xi_1,\cdots,\tau_q+\xi_q,\xi_{q+1},\cdots,\xi_d). \tag{D.8}$$

Thus the derivative of (D.8) factors through the derivative of $\xi\circ J_W$ at $L(\xi_1,\ldots,\xi_d)$, which implies (d); from this (c) follows, because q is the dimension of the orbit through m. If $Y\in P_{\mathcal{F}}$ is defined over an open subset of W, and if we write in the ξ coordinates $Y(\xi)=\sum_i a_i(\xi)\partial/\partial\xi_i$, then, since $Y(\xi)$ is tangent to $\mathcal{O}_{\mathcal{F},\xi}$ by (ii), we deduce from (c), that the functions $a_{q+1},\ldots,a_d$ vanish on $\mathcal{O}_{\mathcal{F},m}$, whence (b) holds.

We finally prove (a). Considering (D.6) and (D.7), a moment's thinking will convince the reader that $W\cap\mathcal{O}_{\mathcal{F},m}$ consists exactly, in the ξ coordinates, of those $(\xi_1,\ldots,\xi_d)$ such that

$$\left(Y^{q+1}_{\xi_{q+1}}\circ\cdots\circ Y^d_{\xi_d}\right)(m)\in\mathcal{O}_{\mathcal{F},m}, \tag{D.9}$$

which accounts for (D.4) where T is the set of $(d-q)$-tuples $(\xi_{q+1},\ldots,\xi_d)$ such that (D.9) holds. To prove that the orbit topology is the product topology on $(-\eta,\eta)^q\times T$ where T is discrete, consider a map F as in (D.1), and pick $\bar t=(\bar t_1,\ldots,\bar t_N)\in\mathrm{dom}(F)$ such that $F(\bar t)\in W$ (hence $F(\bar t)\in W\cap\mathcal{O}_{\mathcal{F},m}$) ; then F is continuous at $\bar t$ for the product topology because, for t close enough to $\bar t$, the values $\xi_{q+1}(F(t)),\ldots,\xi_d(F(t))$ do not depend on t by (b) (moving t_i means following the flow of a vector field in $P_{\mathcal{F}}$, namely the push-forward of X^i through $X^1_{t_1}\circ\cdots\circ X^{i-1}_{t_{i-1}}$) while $\xi_1(F(t)),\ldots,\xi_q(F(t))$ vary continuously with

t according to the continuous dependence on time and initial conditions of solutions to differential equations. Since this is true for all maps F, the orbit topology on $W \cap \mathcal{O}_{\mathcal{F},m}$ is finer than the product topology. To show that it cannot be strictly finer, it is enough to prove that the orbit topology coincides with the Euclidean topology on each set $S_{W,a}$ defined in (D.5), a basis of which consists of the sets $O \times \{a\}$ where O is open in $(-\eta,\eta)^q$. Being open for the product topology, these sets are open the orbit topology as well by what precedes and, since $\mathcal{O}_{\mathcal{F},m}$ is a manifold by (i), each point $(y,a) \in O \times \{a\}$ has, in the orbit topology, a neighborhood $\mathcal{N}_y \subset O \times \{a\}$ which is homeomorphic to an open ball of $\mathbb{R}^q$ *via* some coordinate map. When viewed in these coordinates, the injection $\mathcal{N}_y \to O \times \{a\}$ from the orbit topology to the Euclidean topology is a continuous injective map from an open ball in $\mathbb{R}^q$ into $\mathbb{R}^q$, and therefore it is a homeomorphism onto its image by invariance of the domain. As (y,a) was arbitrary in $O \times \{a\}$, this shows the latter is a union of open sets for the orbit topology, as desired. $\Box$

Consider now the control system :

$$\dot{x} \;=\; f(x,u), \tag{D.10}$$

with state $x \in \mathbb{R}^d$ and control $u \in \mathbb{R}^r$, the function f being smooth on $\mathbb{R}^d \times \mathbb{R}^r$. Let Ω be an open subset of $\mathbb{R}^d \times \mathbb{R}^r$ and, following the notation introduced in section 3, put $\Omega_{\mathbb{R}^d}$ to denote its projection onto the first factor. In the proof of Theorem 6.2, we shall be concerned with the following family of vector fields on $\Omega_{\mathbb{R}^d}$:

$$\mathcal{F}' \;=\; \{\, \delta f_{\alpha_1,\alpha_2}\,,\ \alpha_1,\alpha_2 \text{ feedbacks on } \Omega \,\}\ , \tag{D.11}$$

where feedbacks on Ω were introduced in Definition 3.3 and the notation $\delta f_{\alpha_1,\alpha_2}$ was fixed in (3.5), (3.22).

Since feedbacks are only required to be *continuous*, $\mathcal{F}'$ is a family of continuous *but not necessarily differentiable* vector fields on $\Omega_{\mathbb{R}^d}$ and, though the existence of solutions to differential equations with continuous right-hand side makes it still possible to define the orbit as the collection of endpoints of all concatenated integrations like (D.1), Theorem D.2 does not apply in this case. In fact, if one thinks of the possible non-uniqueness of solutions to the Cauchy problem, a moment's thinking will convince him that the orbits of a family of continuous vector fields have no chance of being topological manifolds in general.

To overcome this difficulty, we will consider instead of $\mathcal{F}'$ the smaller family :

$$\mathcal{F}'' \;=\; \{\, X \in \mathcal{F}'\,,\ X \text{ has a flow} \,\}\ , \tag{D.12}$$

where the sentence "X has a flow" means, as in appendix A, that the Cauchy problem $\dot{x}(t) = X(x(t))$, $x(0) = x_0$, has a unique solution, defined for $|t| < \varepsilon_0$ where ε_0 may depend on x_0, whenever x_0 lies in the domain of definition of X. Let us consider the orbit $\mathcal{O}_{\mathcal{F}'',m}$ of $\mathcal{F}''$ through $m \in \Omega_{\mathbb{R}^d}$, which is still defined as the union of images of all maps (D.1) where $X^j \in \mathcal{F}''$, the domain of each such map F being again a connected open neighborhood $\mathrm{dom}(F)$ of the origin in $\mathbb{R}^N$ by repeated application of Lemma A.1. As before, we define the orbit topology on $\mathcal{O}_{\mathcal{F}'',m}$ to be the finest that makes all the maps (D.1) continuous, and since uniqueness of solutions implies continuous dependence on initial conditions (see [8, chap. V, Theorem 2.1] or Lemma A.1), the orbit topology is again finer than the Euclidean topology. *A priori*, we know very little about $\mathcal{O}_{\mathcal{F}'',m}$ and its orbit topology as Theorem D.2 does not apply. However, Proposition D.5 below will establish that these notions coincide with those arising from the family $\mathcal{F}$ of *smooth* vector fields obtained by setting :

$$\mathcal{F} \;=\; \{\, \delta f_{\alpha_1,\alpha_2}\,,\ \alpha_1,\alpha_2\ \underline{\text{smooth}}\text{ feedbacks on }\Omega\,\}. \tag{D.13}$$

Note that, from the definitions (D.11), (D.12) and (D.13), we obviously have

$$\mathcal{F} \;\subset\; \mathcal{F}'' \;\subset\; \mathcal{F}'\;, \tag{D.14}$$

hence the orbits of these families through a given point obey the same inclusions.

Remark D.4 *It may of course happen that the family $\mathcal{F}'$ is empty because Ω admits no feedback at all. However, if $\mathcal{F}'$ is not empty, then $\mathcal{F}$ is not empty either by Proposition 3.4.*

Proposition D.5 *Suppose that $f : I\!R^d \times I\!R^r \to I\!R^d$ is smooth, and let Ω be an open subset of $I\!R^d \times I\!R^r$. Let $\mathcal{F}''$ be defined by* (D.11)-(D.12).

For any $m \in \Omega_{I\!R^d}$, the orbit $\mathcal{O}_{\mathcal{F}'',m}$ of $\mathcal{F}''$ through m coincides with the orbit through m of the family $\mathcal{F}$ of smooth *vector fields defined by* (D.13)*, and the topology of $\mathcal{O}_{\mathcal{F}'',m}$, as an orbit of $\mathcal{F}$, coincides with its topology as an orbit of $\mathcal{F}''$. In particular, the conclusions of Theorem D.2 hold if we replace $\mathcal{F}$ by $\mathcal{F}''$ and U by $\Omega_{I\!R^d}$.*

Remark D.6 *With a limited amount of extra-work, it is possible to show that the orbits of $\mathcal{F}'$ also coincide with those of $\mathcal{F}$. Hence they turn out to be manifolds despite the possible non-uniqueness of solutions to the Cauchy problem. However,* (D.1) *is no longer convenient to define the orbit topology in this case because the maps F may be multiply-valued when $X^j \in \mathcal{F}'$, and it is simpler to work with the family $\mathcal{F}''$ anyway.*

The proof of the proposition is based on the following lemma.

Lemma D.7 *For $m \in \Omega_{I\!R^d}$ and $X^1,\dots,X^N \in \mathcal{F}''$, let $F : \mathrm{dom}(F) \to \Omega_{I\!R^d}$ be defined by* (D.1)*. Fix $\bar t = (\bar t_1,\dots,\bar t_N) \in \mathrm{dom}(F)$ and set $\overline{m} = F(\bar t)$.*

Then, there is a neighborhood $\mathcal{T}$ of $\bar t$ in $\mathrm{dom}(F)$, with $F(\mathcal{T}) \subset \mathcal{O}_{\mathcal{F},\overline{m}}$, such that $F : \mathcal{T} \to \mathcal{O}_{\mathcal{F},\overline{m}}$ is continuous from the Euclidean topology to the orbit topology.

Assuming the lemma for a while, we first prove the proposition.

Proof of Proposition D.5 We noticed already from (D.14) that the orbit of $\mathcal{F}''$ through m contains the orbit of $\mathcal{F}$ through m. To get the reverse inclusion, consider the map F defined by (D.1) for some vector fields $X^1,\dots,X^N$ belonging to $\mathcal{F}''$. Then, observe from Lemma D.7 that F takes values in a disjoint union of orbits of $\mathcal{F}$, and that it is continuous if each orbit in this union is endowed with the orbit topology. Since $\mathrm{dom}(F)$ is connected, F

takes values in a single orbit, which can be none but $\mathcal{O}_{\mathcal{F},m}$. As F was arbitrary, we conclude that $\mathcal{O}_{\mathcal{F}'',m} \subset \mathcal{O}_{\mathcal{F},m}$ and therefore the two orbits agree as sets. Moreover, since each map F was continuous $\mathrm{dom}(F) \to \mathcal{O}_{\mathcal{F},m}$, the orbit topology of $\mathcal{O}_{\mathcal{F}'',m}$ is by definition finer than the orbit topology of $\mathcal{O}_{\mathcal{F},m}$; but since it is also coarser, by definition of the orbit topology on $\mathcal{O}_{\mathcal{F},m}$, because $\mathcal{F} \subset \mathcal{F}''$, the two topologies in turn agree as desired. $\square$

Proof of Lemma D.7 Theorem D.2 applied to the family $\mathcal{F}$, at the point $\overline{m} = F(\bar{t})$, yields an open neighborhood W of $\overline{m}$ in $\Omega_{\mathbb{R}^d}$ and smooth local coordinates $(\xi_1, \dots, \xi_d) : W \to (-\eta, \eta)^d$ satisfying properties $(iii)-(a)$ to $(iii)-(d)$ of that theorem. For $\varepsilon > 0$ denote by $\mathcal{T}_\varepsilon$ the compact poly-interval :

$$\mathcal{T}_\varepsilon \;=\; \{t = (t_1, \dots, t_N) \in \mathbb{R}^N,\ |t_i - \bar{t}_i| \leq \varepsilon\}\ .$$

By Lemma A.1, F is continuous $\mathrm{dom}(F) \to \Omega_{\mathbb{R}^d}$ and, since $\mathrm{dom}(F)$ is an open neighborhood of $\bar{t}$ in $\mathbb{R}^N$, we can pick $\varepsilon > 0$ such that

$$\mathcal{T}_\varepsilon \subset \mathrm{dom}(F) \quad \text{and} \quad F(\mathcal{T}_\varepsilon) \subset W\ .$$

As $X^1, \dots, X^N$ belong to $\mathcal{F}'' \subset \mathcal{F}'$, we can write

$$X^\ell \;=\; \delta f_{\alpha_1^\ell, \alpha_2^\ell}\ , \quad 1 \leq \ell \leq N$$

for some collection of feedbacks α_1^ℓ, α_2^ℓ on Ω. From Proposition 3.4, there exists for each $(\ell, l) \in \{1, \dots, N\} \times \{1, 2\}$ a sequence of *smooth* feedbacks on Ω, say $(\beta_l^{\ell,k})_{k \in \mathbb{N}}$, converging to α_l^ℓ uniformly on $\Omega_{\mathbb{R}^d}$. Subsequently, we let $Y^{\ell,k}$ denote, for $1 \leq \ell \leq N$ and $k \in \mathbb{N}$, the smooth vector field on $\Omega_{\mathbb{R}^d}$

$$Y^{\ell,k} \;=\; \delta f_{\beta_1^{\ell,k}, \beta_2^{\ell,k}}\ .$$

Clearly $Y^{\ell,k} \in \mathcal{F}$ and, for each ℓ, we have that $Y^{\ell,k}$ converges to X^ℓ as $k \to \infty$, uniformly on compact subsets of $\Omega_{\mathbb{R}^d}$.

Now, pick $j \in \{1, \dots, N\}$ and consider a N-tuple $t^{(j)} \in \mathcal{T}_\varepsilon$ of the form :

$$t^{(j)} = (\bar{t}_1, \dots, \bar{t}_{j-1}, t_j, \dots, t_N)\,, \qquad |t_\ell - \bar{t}_\ell| \leq \varepsilon \ \text{ for } j \leq \ell \leq N.$$

Let also $\mathbf{1}_j$ designate, for simplicity, the N-tuple $(0, \dots, 1, \dots, 0)$ with zero entries except for the j-th one which is 1. Then, for $|\lambda| \leq \varepsilon$, we have that

$$t^{(j)} + \lambda \mathbf{1}_j = (\bar{t}_1, \dots, \bar{t}_{j-1}, \bar{t}_j + \lambda, t_{j+1}, \dots, t_N)\ \in \mathcal{T}_\varepsilon,$$

and a simple computation allows us to rewrite $F(t + \lambda \mathbf{1}_j)$ as :

$$F(t^{(j)} + \lambda \mathbf{1}_j) \;=\; X^1_{\bar{t}_1} \circ \cdots \circ X^{j-1}_{\bar{t}_{j-1}} \circ X^j_\lambda \circ X^{j-1}_{-\bar{t}_{j-1}} \circ \cdots \circ X^1_{-\bar{t}_1}(F(t)).$$

Let us set

$$A_k(\lambda) \;=\; Y^{1,k}_{\bar{t}_1} \circ \cdots \circ Y^{j-1,k}_{\bar{t}_{j-1}} \circ Y^{j,k}_\lambda \circ Y^{j-1,k}_{-\bar{t}_{j-1}} \circ \cdots \circ Y^{1,k}_{-\bar{t}_1}(F(t)).$$

Repeated applications of Lemmas A.1 and A.2 show that, for fixed j and $t^{(j)}$, the map $\lambda \mapsto A_k(\lambda)$ is well-defined $[-\varepsilon, \varepsilon] \to W$ as soon as the integer k is sufficiently large, and moreover that $A_k(\lambda)$ converges to $F(t^{(j)} + \lambda \mathbf{1}_j)$ as $k \to +\infty$, uniformly with respect to $\lambda \in [-\varepsilon, \varepsilon]$. Now, by the characteristic property push forwards, $\lambda \mapsto A_k(\lambda)$ is an integral curve of the smooth vector field

$$Z^k \;=\; \left(Y^{1,k}_{\bar{t}_1} \circ \cdots \circ Y^{j-1,k}_{\bar{t}_{j-1}}\right)_\star Y^{j,k} \;,$$

which is defined on a neighborhood of $\{F(t^{(j)} + \lambda \mathbf{1}_j);\ |\lambda| \leq \varepsilon\}$ in W. Since $Z^k \in P_{\mathcal{F}}$ (*cf* equation (D.3)), it follows from point $(iii) - (b)$ of Theorem D.2 that, for k large enough,

$$\xi_i \circ A_k(\lambda) \;=\; \xi_i \circ A_k(0) \;,\quad \forall \lambda \in [-\varepsilon, \varepsilon],\ i \in \{q+1, \ldots, d\}.$$

It is clear from the definition that $A_k(0) = F(t^{(j)})$; hence, using the continuity of ξ_i and taking, in the above equation, the limit as $k \to +\infty$, we get

$$\xi_i \circ F(t^{(j)} + \lambda \mathbf{1}_j) = \xi_i \circ F(t^{(j)}),\ \forall \lambda \in [-\varepsilon, \varepsilon],\ i \in \{q+1, \ldots, d\}. \tag{D.15}$$

Since $\xi_{q+1} \circ F(\bar{t}) = \cdots = \xi_d \circ F(\bar{t}) = 0$ by definition of W, successive applications of (D.15) for $j = N, \ldots, 1$ lead us to the conclusion that

$$\xi_{q+1} \circ F(t) = \cdots = \xi_d \circ F(t) = 0,\quad \forall t \in \mathcal{T}_\varepsilon \;. \tag{D.16}$$

Equation (D.16) means that, in the ξ-coordinates, $F(\mathcal{T}_\varepsilon) \subset (-\eta, \eta)^q \times \{0\}$. Hence, from the local description of the orbits in (D.4) (where m is to be replaced by $\overline{m}$), we deduce that $F(\mathcal{T}_\varepsilon) \subset \mathcal{O}_{\mathcal{F},\overline{m}}$. Actually, with the notations of (D.5), we even get the stronger conclusion that

$$F(\mathcal{T}_\varepsilon) \;\subset\; S_{W,0}$$

which achieves the proof of the lemma, with $\mathcal{T} = \mathcal{T}_\varepsilon$, because the orbit topology on $S_{W,0}$ is the Euclidean topology by Remark D.3. □

E An averaging lemma

The following averaging lemma for continuous vector fields is less classical than in the locally Lipschitz case, where uniqueness of solutions to the Cauchy problem prevails.

Lemma E.1 *Let $t_1 < t_2$ be real numbers and $(X^{1,\ell})_{\ell \in \mathbb{N}}$, $(X^{2,\ell})_{\ell \in \mathbb{N}}$, be two sequences of continuous time-dependent vector fields $[t_1, t_2] \times \mathbb{R}^d \to \mathbb{R}^d$, uniformly bounded with respect to ℓ, that converge uniformly on compact subsets of $[t_1, t_2] \times \mathbb{R}^d$ to some vector fields X^1 and X^2 respectively. Denoting by $L = t_2 - t_1$ the length of the time interval, define, for each $\ell \in \mathbb{N}$, the "average" vector field $G_\ell : [t_1, t_2] \times \mathbb{R}^d \to \mathbb{R}^d$ by :*

$$\begin{array}{l} t \in [t_1 + \frac{j}{\ell} L\,,\, t_1 + \frac{2j+1}{2\ell} L) \;\Rightarrow\; G_\ell(t,x) \;=\; X^{1,\ell}(t,x)\;, \\ t \in [t_1 + \frac{2j+1}{2\ell} L\,,\, t_1 + \frac{j+1}{\ell} L) \;\Rightarrow\; G_\ell(t,x) \;=\; X^{2,\ell}(t,x) \quad , \end{array} \tag{E.1}$$

for $j \in \{0, \ldots, \ell-1\}$ *and, say,* $G_\ell(t_2, x) = X^{2,\ell}(t_2, x)$ *for definiteness.*
Let $\gamma_\ell : [t_1, t_2] \to I\!R^d$ *be a solution to*

$$\gamma_\ell(t) - \bar{x} = \int_{t_1}^{t} G_\ell(\tau, \gamma_\ell(\tau)) \mathrm{d}\tau . \tag{E.2}$$

Then the sequence (γ_ℓ) *is compact in* $C^0([t_1, t_2], I\!R^d)$*, and every accumulation point* γ_∞ *is a solution to*

$$\gamma_\infty(t) - \bar{x} = \frac{1}{2} \int_{t_1}^{t} \left(X^1(\tau, \gamma_\infty(\tau)) + X^2(\tau, \gamma_\infty(\tau)) \right) \mathrm{d}\tau . \tag{E.3}$$

Proof. Let

$$M = \sup_{t,x,i,\ell} \|X^{i,\ell}(t, x)\| . \tag{E.4}$$

From (E.1)-(E.2), it is clear that M is a Lipschitz constant for γ_ℓ, regardless of ℓ. In particular $\gamma_\ell(t)$ stays in a fixed compact ball B of radius ML, and the family (γ_ℓ) is equicontinuous. From Ascoli-Arzela's theorem this implies compactness of the sequence (γ_ℓ) in the uniform topology on $[t_1, t_2]$.

Rewrite (E.2) as

$$\begin{aligned} \gamma_\ell(t) - \bar{x} = & \int_{t_1}^{t} \left(G_\ell(\tau, \gamma_\ell(\tau)) - \frac{X^{1,\ell}(\tau, \gamma_\ell(\tau)) + X^{2,\ell}(\tau, \gamma_\ell(\tau))}{2} \right) \mathrm{d}\tau \\ & + \int_{t_1}^{t} \left(\frac{X^{1,\ell}(\tau, \gamma_\ell(\tau)) + X^{2,\ell}(\tau, \gamma_\ell(\tau))}{2} - \frac{X^{1}(\tau, \gamma_\ell(\tau)) + X^{2}(\tau, \gamma_\ell(\tau))}{2} \right) \mathrm{d}\tau \\ & + \int_{t_1}^{t} \frac{X^{1}(\tau, \gamma_\ell(\tau)) + X^{2}(\tau, \gamma_\ell(\tau))}{2} \mathrm{d}\tau . \end{aligned} \tag{E.5}$$

By the uniform convergence of $X^{i,\ell}$ to X^i, it will clearly follow that any accumulation point γ_∞ of (γ_ℓ) satisfies (E.3) if only we can show that the first integral in the right-hand side of (E.5) converges to zero as $\ell \to \infty$.

To prove this, we compute, from the definition of G_ℓ :

$$\begin{aligned} & \int_{t_1 + \frac{j}{\ell} L}^{t_1 + \frac{j+1}{\ell} L} \left(G_\ell(\tau, \gamma_\ell(\tau)) - \frac{X^{1,\ell}(\tau, \gamma_\ell(\tau)) + X^{2,\ell}(\tau, \gamma_\ell(\tau))}{2} \right) \mathrm{d}\tau \\ = & \int_{t_1 + \frac{j}{\ell} L}^{t_1 + \frac{2j+1}{2\ell} L} \frac{X^{1,\ell}(\tau, \gamma_\ell(\tau)) - X^{2,\ell}(\tau, \gamma_\ell(\tau))}{2} \mathrm{d}\tau \\ & \quad - \int_{t_1 + \frac{2j+1}{2\ell} L}^{t_1 + \frac{j+1}{\ell} L} \frac{X^{1,\ell}(\tau, \gamma_\ell(\tau)) - X^{2,\ell}(\tau, \gamma_\ell(\tau))}{2} \mathrm{d}\tau \\ = & \int_{t_1 + \frac{j}{\ell} L}^{t_1 + \frac{2j+1}{2\ell} L} \left(\Delta_\ell(\tau, \gamma_\ell(\tau)) - \Delta_\ell(\tau + \tfrac{L}{2\ell}, \gamma_\ell(\tau + \tfrac{L}{2\ell})) \right) \mathrm{d}\tau \end{aligned} \tag{E.6}$$

with $\Delta_\ell = \frac{1}{2}(X^{1,\ell} - X^{2,\ell})$. On the compact set $[t_1, t_2] \times B$, the vector field Δ_ℓ is uniformly continuous with a modulus of continuity that does not depend on ℓ ; consequently, by the uniform Lipschitz property of γ_ℓ, we see for arbitrary $\varepsilon > 0$ that the norm of the last integral is less that $\varepsilon/2\ell$ as soon as ℓ is large enough, independently of j.

Now, the first integral in (E.5) can be decomposed into a sum of at most ℓ integrals like these we just studied plus an integral over an interval of length smaller that $1/\ell$. Since the norm of the integrand is bounded by $2M$, the norm of the last term is less than $2M/\ell$. Summing over j, the above estimates tell us that, for $t \in [t_1, t_2]$ and for ℓ is large enough,

$$\int_{t_1}^{t} \left(G_\ell(\tau, \gamma_\ell(\tau)) - \frac{X^1(\tau, \gamma_\ell(\tau)) + X^2(\tau, \gamma_\ell(\tau))}{2} \right) \mathrm{d}\tau \leq \frac{\varepsilon}{2} + \frac{2M}{\ell}.$$

This achieves the proof since $\varepsilon > 0$ was arbitrary. $\square$

F Key lemmas to Theorem 6.1

The following two lemmas will be applied recursively in the proof of Theorem 6.2 to obtain the forms (6.8), (6.5), and (6.9). Although these lemmas team up into a single result in the above-mentioned proof, they have been stated here separately for the sake of clarity.

We will consider two control systems with state in $\mathbb{R}^d$ and control in $\mathbb{R}^r$. Expanded in coordinates, the first system reads

$$\begin{array}{rcl} \dot{x}_1 & = & f_1(x_1, \ldots, x_d, x_{d+1}, \ldots, x_{d+r}) \\ & \vdots & \\ \dot{x}_d & = & f_d(x_1, \ldots, x_d, x_{d+1}, \ldots, x_{d+r}) \ , \end{array} \tag{F.1}$$

with state variable $(x_1, \ldots, x_d)$ and control variable $(x_{d+1}, \ldots, x_{d+r}) \in \mathbb{R}^r$, the functions $f_1, \cdots, f_d$ being smooth $\mathbb{R}^{d+r} \to \mathbb{R}$. The second system has state variable $(z_1, \ldots, z_d)$ and control variable $(z_{d+1}, \ldots, z_{d+r}) \in \mathbb{R}^r$, and it assumes the special form :

$$\begin{array}{rcl} \dot{z}_1 & = & g_1(z_1, \ldots, z_d) \\ & \vdots & \\ \dot{z}_{d-s} & = & g_{d-s}(z_1, \ldots, z_d) \\ \dot{z}_{d-s+1} & = & z_{d+1} \\ & \vdots & \\ \dot{z}_d & = & z_{d+s} \ , \end{array} \tag{F.2}$$

where $0 < s \leq d$ and $s \leq r$ while $g_1, \cdots, g_{d-s}$ are again smooth $\mathbb{R}^d \to \mathbb{R}$. Nothing prevents us here from having $s < r$, in which case some of the controls do not enter the equation. It will be convenient to use the aggregate notations

$$\begin{array}{ll} X \triangleq (x_1, \ldots, x_d)\,, & U \triangleq (x_{d+1}, \ldots, x_{d+r})\,, \\ Z \triangleq (z_1, \ldots, z_d)\,, & V \triangleq (z_{d+1}, \ldots, z_{d+r})\,, \end{array}$$

and to further split Z into (Z^1, Z^2) with

$$Z^1 \triangleq (z_1, \ldots, z_{d-s}), \quad Z^2 \triangleq (z_{d-s+1}, \ldots, z_d), \tag{F.3}$$

so as to write (F.1) in the form

$$\dot{X} = f(X, U) \tag{F.4}$$

and (F.2) as

$$\begin{aligned} \dot{Z}^1 &= g^1(Z^1, Z^2) \\ \dot{Z}^2 &= J_r^s V, \end{aligned} \tag{F.5}$$

with J_r^s the $s \times r$ matrix, defined in (4.8), that selects the first s entries of a vector.

Lemma F.1 *Let d, r and s be strictly positive integers with $s \leq d$ and $s \leq r$. Suppose, for some $\varepsilon > 0$, that*

$$\varphi : (-\varepsilon, \varepsilon)^{d+r} \to \mathbb{R}^{d+r}$$

is a homeomorphism onto its image, with inverse ψ, that conjugates system (F.4) *to system* (F.5). *Then, there exists $0 < \varepsilon' < \varepsilon$ and a* smooth *local change of coordinates around $0 \in \mathbb{R}^d$:*

$$\theta : (-\varepsilon', \varepsilon')^d \to \theta\big((-\varepsilon', \varepsilon')^d\big) \subset (-\varepsilon, \varepsilon)^d$$

that fixes the origin and is such that, in the new coordinates $\widetilde{X} = \theta^{-1}(X)$, both the system (F.4) *and the conjugating homeomorphism $\widetilde{\varphi} = \varphi \circ (\theta \times \mathrm{id})$ assume a block triangular structure with respect to the partition $\widetilde{X} = (\widetilde{X}^1, \widetilde{X}^2)$, where $\widetilde{X}^1 \triangleq (\widetilde{x}_1, \ldots, \widetilde{x}_{d-s})$ and $\widetilde{X}^2 \triangleq (\widetilde{x}_{d-s+1}, \ldots, \widetilde{x}_d)$; that is to say, on $(-\varepsilon', \varepsilon')^{d+r}$, we have that*

- *system* (F.1) *reads :*

$$\begin{aligned} \dot{\widetilde{X}}^1 &= \widetilde{f}^1(\widetilde{X}^1, \widetilde{X}^2) \\ \dot{\widetilde{X}}^2 &= \widetilde{f}^2(\widetilde{X}^1, \widetilde{X}^2, U), \end{aligned} \tag{F.6}$$

- *On their respective domains of definition, the homeomorphism $\widetilde{\varphi}$ and its inverse $\widetilde{\psi} = (\theta^{-1} \times \mathrm{id}) \circ \psi$ read :*

$$\begin{aligned} Z^1 &= \widetilde{\varphi}^1(\widetilde{X}^1) & \widetilde{X}^1 &= \widetilde{\psi}_1(Z^1) \\ Z^2 &= \widetilde{\varphi}^2(\widetilde{X}^1, \widetilde{X}^2) & \widetilde{X}^2 &= \widetilde{\psi}_2(Z^1, Z^2) \\ V &= \widetilde{\varphi}^3(\widetilde{X}^1, \widetilde{X}^2, U) & U &= \widetilde{\psi}_3(Z^1, Z^2, V). \end{aligned} \tag{F.7}$$

Lemma F.2 *Let*

$$\widetilde{\varphi} : (-\varepsilon', \varepsilon')^{d+r} \to \mathbb{R}^{d+r}$$

be a homeomorphism onto its image, having the block triangular structure displayed in (F.7), *and assume that it conjugates the smooth system* (F.6) *to the smooth system* (F.5). *Necessarily then, $\widetilde{\varphi}$ has the following properties :*

1. *The map* $\widetilde{\varphi}^2$ *is continuously differentiable with respect to its second argument* $\widetilde{X}^2$, *and* $\dfrac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(0,0)$ *is invertible.*
2. *On some neighborhood of* $0 \in \mathbb{R}^{d+r}$ *included in* $(-\varepsilon', \varepsilon')^{d+r}$, *one has :*

$$\begin{aligned} \widetilde{f}^2(\widetilde{X}^1, \widetilde{X}^2, U) \quad &= \\ \widetilde{f}^2(\widetilde{X}^1, \widetilde{X}^2, 0) \quad &+ \quad \left(\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1, \widetilde{X}^2)\right)^{-1} J_r^s \left(\widetilde{\varphi}^3(\widetilde{X}^1, \widetilde{X}^2, U) - \widetilde{\varphi}^3(\widetilde{X}^1, \widetilde{X}^2, 0)\right) \end{aligned} \tag{F.8}$$

3. *On some neighborhood of* $0 \in \mathbb{R}^d$ *included in* $(-\varepsilon', \varepsilon')^d$, *the partial homeomorphism*

$$(\,\widetilde{X}^1\,,\,\widetilde{X}^2\,) \quad \mapsto \quad (\,\widetilde{\varphi}^1(\widetilde{X}^1)\,,\,\widetilde{\varphi}^2(\widetilde{X}^1, \widetilde{X}^2)\,) \tag{F.9}$$

conjugates the control system

$$\dot{\widetilde{X}}^1 = \widetilde{f}^1(\widetilde{X}^1, \widetilde{X}^2), \tag{F.10}$$

with state $\widetilde{X}^1$ *and control* $\widetilde{X}^2$, *to the control system*

$$\dot{Z}^1 = g^1(Z^1, Z^2) \tag{F.11}$$

with state Z^1 *and input* Z^2.

Note that (F.10) and (F.11) are reduced systems from (F.6) and (F.5).

Proof of Lemma F.1. Since the homeomorphism φ conjugates (F.4) to (F.5), we know, by Proposition 3.7, that φ and ψ split component-wise into :

$$\begin{array}{rclcrcl} Z &=& \varphi_{\mathrm{I}}(X) & \qquad & X &=& \psi_{\mathrm{I}}(Z) \\ V &=& \varphi_{\mathrm{II}}(X, U) & \qquad & U &=& \psi_{\mathrm{II}}(Z, V)\ . \end{array} \tag{F.12}$$

Consider the map $f:\ (-\varepsilon, \varepsilon)^{d+r} \to \mathbb{R}^d$ given in (F.4), and let us define $g : \varphi((-\varepsilon, \varepsilon)^{d+r}) \to \mathbb{R}^d$ analogously from (F.5), namely g is the concatenated map whose first $d-s$ components are given by $g^1(Z)$ and whose last s components are given by $J_r^s V$. Define two families of continuous vector fields $\mathcal{F}'$ and $\mathcal{G}'$, on $(-\varepsilon, \varepsilon)^d$ and $\varphi_{\mathrm{I}}((-\varepsilon, \varepsilon)^d)$ respectively, by the following formulas (compare (D.11)) :

$$\mathcal{F}' \quad = \quad \{\,\delta f_{\alpha_1,\alpha_2}\,;\alpha_1, \alpha_2 \text{ feedbacks on } (-\varepsilon, \varepsilon)^{d+r}\,\}\ , \tag{F.13}$$

$$\mathcal{G}' \quad = \quad \{\,\delta g_{\beta_1,\beta_2}\,;\beta_1, \beta_2 \text{ feedbacks on } \varphi\big((-\varepsilon, \varepsilon)^{d+r}\big)\,\}\ . \tag{F.14}$$

Applying Proposition 3.11 twice, first to $\chi = \varphi$ and then to $\chi = \psi$, we see that each integral curve of a vector field in $\mathcal{F}'$ is mapped by φ_{I} to some integral curve of a vector field in $\mathcal{G}'$ and *vice-versa* upon replacing φ_{I} by ψ_{I}. This shows in particular that uniqueness of solutions to the Cauchy problem associated to vector fields is preserved, i.e. if we define the families of vector fields (compare (D.12)) :

$$\mathcal{F}'' = \{ Y \in \mathcal{F}' ,\ Y \text{ has a flow} \} , \tag{F.15}$$
$$\mathcal{G}'' = \{ Y \in \mathcal{G}' ,\ Y \text{ has a flow} \} , \tag{F.16}$$

we also have that each integral curve of a vector field in $\mathcal{F}''$ is mapped by φ_{I} to an integral curve of a vector field in $\mathcal{G}''$ and *vice-versa* upon replacing φ_{I} by ψ_{I}. By concatenation, using Proposition D.5, it follows that

$$\left.\begin{array}{l} \textit{for any } X \in (-\varepsilon,\varepsilon)^d,\ \varphi_{\mathrm{I}} \textit{ defines a homeomorphism,} \\ \textit{for the orbit topologies, from the orbit of } \mathcal{F}'' \textit{ through } \mathrm{X} \\ \textit{onto the orbit of } \mathcal{G}'' \textit{ through } \varphi_{\mathrm{I}}(X), \end{array}\right\} \tag{F.17}$$

where the orbit topology as described in Proposition D.5 (by definition the restriction of φ_{I} is bi-continuous for the topologies induced by the ambient space; bi-continuity for the orbit topologies requires the description of these topologies as given in Proposition D.5).

Now, the vector fields $\delta g_{\beta_1,\beta_2}$ appearing in (F.14) inherit from the structure of g, displayed in (F.5), the following particular form :

$$\delta g_{\beta_1,\beta_2}(Z) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta_{1,1}(Z) - \beta_{2,1}(Z) \\ \vdots \\ \beta_{1,s}(Z) - \beta_{2,s}(Z) \end{pmatrix} , \tag{F.18}$$

where $\beta_{i,1}, \ldots, \beta_{i,s}$ designate, for $i = 1, 2$, the first s component of the feedback β_i. This will allow for us to describe explicitly the orbits of $\mathcal{G}''$, namely :

$$\left.\begin{array}{l} \textit{the orbit of } \mathcal{G}'' \textit{ through } Z_0 = (c_1, \ldots, c_d) \\ \textit{is the connected component containing } Z_0 \textit{ of the set} \\ \{Z \in \varphi_{\mathrm{I}}\left((-\varepsilon,\varepsilon)^d\right) ,\ z_1 = c_1, \ldots, z_{d-s} = c_{d-s}\}. \end{array}\right\} \tag{F.19}$$

Indeed, the orbit in question is contained in this set, because it is connected, and because all the vector fields in $\mathcal{G}''$ have their first $d - s$ components equal to zero by (F.18).

To prove the reverse inclusion, it is enough to show that the orbit of $\mathcal{G}''$ through Z_0, denoted hereafter by $\mathcal{O}_{\mathcal{G}'',Z_0}$, contains all the points sufficiently close to Z_0 having the same first $d-s$ coordinates as Z_0. Indeed, since Z_0 was arbitrary, this will imply that the connected component defined by (F.19) splits into a disjoint union of open orbits hence consists of a single one by connectedness. That is to say, putting $Z_0 = (Z_0^1, Z_0^2)$ according to (F.3), F.19 will follow from the existence of a $\rho > 0$ such that

$$\{Z_0^1\} \times B(Z_0^2, \rho) = B(Z_0, \rho) \cap \mathcal{O}_{\mathcal{G}'',Z_0}. \tag{F.20}$$

Now, it follows from Remark D.3 that, for sufficiently small ρ, each connected component of $B(Z_0,\rho)\cap\mathcal{O}_{\mathcal{G}'',Z_0}$ is an embedded sub-manifold of $B(Z_0,\rho)$. Then, the connected component of $B(Z_0,\rho)\cap\mathcal{O}_{\mathcal{G}'',Z_0}$ containing Z_0 is, by inclusion, an embedded sub-manifold of the linear manifold $\{Z_0^1\}\times B(Z_0^2,\rho)$. In particular, since no strict sub-manifold can be densely embedded in a given manifold, we see that (F.20) will hold is only we can prove that

$$\begin{array}{l}\textit{The connected component containing } Z_0 \textit{ of } B(Z_0,\rho)\cap\mathcal{O}_{\mathcal{G}'',Z_0}\\ \textit{is dense in } \{Z_0^1\}\times B(Z_0^2,\rho) \textit{ for the Euclidean topology.}\end{array}\tag{F.21}$$

To prove (F.21), pick V_0 such that $(Z_0,V_0)\in\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$ and observe, since the latter is an open set, that shrinking ρ further, if necessary, allows us to assume $\overline{B}(Z_0,\rho)\times\overline{B}(V_0,\rho)\subset\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$. *We claim that any continuous map $\overline{B}(Z_0,\rho)\to\overline{B}(V_0,\rho)$ extends to a feedback on $\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$.* Indeed, in view of the one-to-one correspondence $\beta\to\psi\blacksquare\beta$ between feedbacks on $\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$ and feedbacks on $(-\varepsilon,\varepsilon)^{d+r}$ (*cf* the discussion leading to (3.20)-(3.21)), it is enough to prove that every continuous map $\psi_{\mathrm{I}}\big(\overline{B}(Z_0,\rho)\big)\to(-\varepsilon,\varepsilon)^r$ extends to a continuous map $(-\varepsilon,\varepsilon)^d\to(-\varepsilon,\varepsilon)^r$, and this in turn follows from the Tietze extension theorem since $\psi_{\mathrm{I}}\big(\overline{B}(Z_0,\rho)\big)$ is closed in $(-\varepsilon,\varepsilon)^d$ and since $(-\varepsilon,\varepsilon)^r$ is a poly-interval. *This proves the claim.*

From the claim, it follows that the restriction to $\overline{B}(Z_0,\rho)$ of the $I\!R^s$-valued vector field $J_r^s(\beta_1(Z)-\beta_2(Z))$, accounting for the lower half of the right-hand side in (F.18), can be assigned *arbitrarily*, by choosing adequately the feedbacks β_1 and β_2, among *continuous* vector fields $\overline{B}(Z_0,\rho)\to\overline{B}(0,\rho)$ (take β_2 to extend the constant map V_0 on $\overline{B}(Z_0,\rho)$). Of course, the corresponding vector field $\delta g_{\beta_1,\beta_2}$ in (F.18) belongs to $\mathcal{G}'$ but not necessarily to $\mathcal{G}''$ since continuous vector fields need not have a flow. However, since $\delta g_{\beta_1,\beta_2}$ has a flow at least when β_1 and β_2 are smooth, we deduce from Proposition 3.4 that the restriction to $\overline{B}(Z_0,\rho)$ of the vector fields in $\mathcal{G}''$ are of the form $\{0\}\times Y$, where Y ranges over a uniformly dense subset Υ of all $I\!R^s$-valued continuous maps $\overline{B}(Z_0,\rho)\to\overline{B}(0,\rho)$. Now, every point in $B(Z_0^2,\rho)$ can be attained from Z_0^2 upon integrating, *within* $B(Z_0^2,\rho)$, a constant vector field of arbitrary small norm. By Lemma A.2 applied with $\mathcal{U}=B(Z_0^2,\rho)$ and $K=\{Z_0^2\}$, the corresponding trajectory can be approximated uniformly by integral curves *that remain in* $B(Z_0^2,\rho)$ of vector fields in Υ. Therefore, every point in $\{z_0^1\}\times B(Z_0^2,\rho)$ is the limit of endpoints of integral curves of $\mathcal{G}''$ *that remain in* $\{z_0^1\}\times B(Z_0^2,\rho)$, which proves (F.21) and thus (F.19). In particular, the orbits of $\mathcal{G}''$ are *embedded* sub-manifolds in $\varphi_{\mathrm{I}}\big((-\varepsilon,\varepsilon)^d\big)$.

Next, we turn to the orbits of $\mathcal{F}''$, and we designate by $\mathcal{O}_{\mathcal{F}'',p}$ the orbit of $\mathcal{F}''$ in $]-\varepsilon,\varepsilon[^d$ through the point p. On the one hand, Proposition D.5 and Theorem D.2 show that $\mathcal{O}_{\mathcal{F}'',p}$ is a smooth immersed sub-manifold of $]-\varepsilon,\varepsilon[^d$. On the other hand, by (F.17), this immersed sub-manifold is sent homeomorphically by φ_{I}, both for the orbit topology and the ambient topology, onto $\mathcal{O}_{\mathcal{G}'',\varphi_{\mathrm{I}}(p)}$ which is a smooth *embedded* s-dimensional sub-manifold of $\varphi_{\mathrm{I}}\big((-\varepsilon,\varepsilon)^d\big)$, as we saw from (F.19). This entails that all orbits of $\mathcal{F}''$ in $]-\varepsilon,\varepsilon[^d$ are *embedded* sub-manifolds of dimension s. Consequently, still from Proposition D.5 and Theorem D.2, there are coordinates $(\xi_1,\dots,\xi_d)$ defined on an open neighborhood W_0 of the origin in $]-\varepsilon,\varepsilon[^d$ —this neighborhood may be assumed to be of the form $\{(\xi_1,\dots,\xi_d),\ |\xi_i|<\varepsilon'\}$ —

such that, in these coordinates,

$$W_0 \cap \mathcal{O}_{\mathcal{F}'',0} \;=\; \{\, (\xi_1,\dots,\xi_d), \text{ with } (\xi_{s+1},\dots,\xi_d) \in T \,\}\ ,$$

with T a subset of $]-\varepsilon',\varepsilon'[^{d-s}$ containing $(0,\dots,0)$, the tangent space to $W_0 \cap \mathcal{O}_{\mathcal{F}'',0}$ at each of its points being *spanned* by $\partial/\partial\xi_1,\dots,\partial/\partial\xi_s$, while at any point $p \in W_0$ the vector fields $\partial/\partial\xi_1,\dots,\partial/\partial\xi_s$ *belong* to the tangent space of $\mathcal{O}_{\mathcal{F}'',p}$. But since we saw that *all* orbits are smooth sub-manifolds of dimension s, these vector fields actually *span* the tangent space to the orbit at every point. Hence all the vector fields $\delta f_{\alpha_1,\alpha_2}$ in $\mathcal{F}''$ have their last $d-s$ components equal to zero on W_0 in the ξ coordinates, and this holds in particular when α_1, α_2 range over all constant feedbacks $(-\varepsilon,\varepsilon)^d \to (-\varepsilon,\varepsilon)^r$. This implies, by the very definition of $\delta f_{\alpha_1,\alpha_2}$, that $(\dot\xi_{s+1},\dots,\dot\xi_d)$ — as computed from (F.4) upon performing the change of variable $X \mapsto (\xi_1,\dots,\xi_d)$ — does not depend on the control variable U. Choose for $\widetilde{X}$ the ξ coordinates arranged in reverse order, and let $\widetilde{f}$ be the analog of f in the new coordinates $(\widetilde{X},U)$. Then the first $d-s$ components of $\widetilde{f}$ do not depend on U so that (F.6) holds. Moreover, if $\widetilde{\varphi}$ denotes the new homeomorphism that conjugates (F.6) to (F.5) over $(-\varepsilon,\varepsilon)^{d+r}$, $\widetilde{\varphi}((-\varepsilon,\varepsilon)^{d+r})$, and if $\widetilde{\psi}$ denotes its inverse, it follows from (F.17) and the above characterization of the orbits that $\widetilde{\varphi}_{\mathrm{I}}$ maps the sets where $\tilde{x}_1,\dots,\tilde{x}_{d-s}$ are constant to those where $z_1,\dots,z_{d-s}$ are constant, thus the functions $\widetilde{\varphi}_1,\dots,\widetilde{\varphi}_{d-s}$ and $\widetilde{\psi}_1,\dots,\widetilde{\psi}_{d-s}$ depend only on their $d-s$ first arguments whence (F.7) follows. $\square$

Proof of Lemma F.2 We use again the concatenated notation $\widetilde{\varphi}_{\mathrm{I}} = (\widetilde{\varphi}^1,\widetilde{\varphi}^2)$, $\widetilde{\psi}_{\mathrm{I}} = (\widetilde{\psi}^1,\widetilde{\psi}^2)$, these partial homeomorphisms being inverse of each other. Let $(Z_0,V_0) \in \widetilde{\varphi}((-\varepsilon',\varepsilon')^{d+r})$ and ε'' be so small that the product neighborhood $(Z_0,V_0)+(-\varepsilon'',\varepsilon'')^{d+r}$ lies entirely within $\widetilde{\varphi}((-\varepsilon',\varepsilon')^{d+r})$. The restriction to $(Z_0,V_0)+(-\varepsilon'',\varepsilon'')^{d+r}$ of $\widetilde{\psi}$ conjugates (F.5) to (F.6). Consequently, for any $\overline{V} \in (-\varepsilon'',\varepsilon'')^r$, we may apply Proposition 3.11 to this restriction and to the constant feedbacks $\alpha_1(Z) = V_0 + \overline{V}$ and $\alpha_2(Z) = V_0$; this yields that $\widetilde{\psi}_{\mathrm{I}}$, given by

$$(Z^1,Z^2) \;\mapsto\; (\widetilde{X}^1,\widetilde{X}^2) \;=\; (\widetilde{\psi}^1(Z^1),\widetilde{\psi}^2(Z^1,Z^2)),$$

maps every solution of

$$\dot{Z}^1 \;=\; 0 \ , \qquad \dot{Z}^2 \;=\; J_r^s\overline{V} \tag{F.22}$$

that remains in $Z_0 + (-\varepsilon'',\varepsilon'')^d$ to a solution of

$$\begin{aligned}\dot{\widetilde{X}}^1 = 0\ , \quad \dot{\widetilde{X}}^2 &= \widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),V_0+\overline{V}))\\ &\quad - \widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),V_0))\end{aligned} \tag{F.23}$$

that remains in $\widetilde{\psi}_{\mathrm{I}}(Z_0 + (-\varepsilon'',\varepsilon'')^d)$, and *vice versa* upon applying Proposition 3.11 in the other direction.

Integrating (F.22) explicitly with initial condition $Z(0) = Z_0$, we get that

$$t \;\mapsto\; \begin{pmatrix} \widetilde{\psi}^1(Z_0^1) \\ \widetilde{\psi}^2(Z_0^1, Z_0^2 \,+\, tJ_r^s\overline{V}) \end{pmatrix}$$

solves (F.23) for sufficiently small t, hence $\widetilde{\psi}^2(Z^1,Z^2)$ is differentiable at Z_0 with respect to its second argument in the direction $J_r^s\overline{V}$, with directional derivative

$$\begin{aligned}\frac{\partial\widetilde{\psi}^2}{\partial Z^2}(Z_0^1,Z_0^2)\,J_r^s\overline{V} &= \widetilde{f}^2(\widetilde{\psi}^1(Z_0^1),\widetilde{\psi}^2(Z_0^1,Z_0^2),\widetilde{\psi}^3(Z_0^1,Z_0^2,V_0+\overline{V}))\\ &\quad -\widetilde{f}^2(\widetilde{\psi}^1(Z_0^1),\widetilde{\psi}^2(Z_0^1,Z_0^2),\widetilde{\psi}^3(Z_0^1,Z_0^2,V_0))\,. \end{aligned}\tag{F.24}$$

In particular, since Z_0 can be any member of $\widetilde{\varphi}_{\mathrm{I}}((-\varepsilon',\varepsilon')^d)$ while $J_r^s\overline{V}$ can be assigned arbitrarily in $(-\varepsilon'',\varepsilon'')^s$, we conclude that $\partial\widetilde{\psi}^2/\partial Z^2(Z^1,Z^2)$ exists and is continuous since this holds for the partial derivatives. Next we prove that $\partial\widetilde{\psi}^2/\partial Z^2$ is invertible at every point by showing that its kernel reduces to zero. In fact, if the left-hand side of (F.24) vanishes, so does the right-hand side which is also the value of the right-hand side of (F.23) for $\widetilde{X}=\widetilde{\psi}_{\mathrm{I}}(Z_0)$. Therefore the constant map $t\mapsto\widetilde{\psi}_{\mathrm{I}}(Z_0)$ is a solution to (F.23) over a suitable time interval, and by conjugation the constant map $t\mapsto Z_0$ is a solution to (F.22) over that time interval which clearly entails $J_r^s\overline{V}=0$, as desired. Now, since $\partial\widetilde{\psi}^2/\partial Z^2$ is invertible at every $(Z^1,Z^2)\in\widetilde{\varphi}_{\mathrm{I}}((-\varepsilon',\varepsilon')^d)$, the triangular structure of (F.7) and the inverse function theorem together imply that

$$\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1,\widetilde{X}^2)=\left(\frac{\partial\widetilde{\psi}^2}{\partial Z^2}(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2))\right)^{-1}\tag{F.25}$$

continuously exists and is invertible for $(\widetilde{X}^1,\widetilde{X}^2)\in(-\varepsilon',\varepsilon')^d$. This proves point 1.

Let us turn to point 2. Select an open neighborhood $\mathcal{W}$ of 0 having compact closure in $(-\varepsilon',\varepsilon')^d$, so there is $\eta>0$ such that $\widetilde{\varphi}(\widetilde{X},0)+(-\eta,\eta)^{d+r}$ is included in $\widetilde{\varphi}((-\varepsilon',\varepsilon')^{d+r})$ whenever $\widetilde{X}\in\mathcal{W}$. If $\overline{V}\in(-\eta,\eta)^r$, we can apply (F.24) to $(Z_0,V_0)=\widetilde{\varphi}(\widetilde{X},0)$ with $\widetilde{X}\in\mathcal{W}$, and we obtain in view of (F.25) :

$$\begin{aligned}\left(\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1,\widetilde{X}^2)\right)^{-1} J_r^s\overline{V} &= -\,\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,0)\\ &\quad +\widetilde{f}^2\left(\widetilde{X}^1,\widetilde{X}^2,\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),\widetilde{\varphi}^3(\widetilde{X}^1,\widetilde{X}^2,0)+\overline{V})\right)\,.\end{aligned}\tag{F.26}$$

Set

$$U=\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),\widetilde{\varphi}^3(\widetilde{X}^1,\widetilde{X}^2,0)+\overline{V})\tag{F.27}$$

and observe that $(\widetilde{X},\overline{V})\mapsto(\widetilde{X},U)=\widetilde{\psi}(\widetilde{\varphi}(\widetilde{X},0)+(0,\overline{V}))$ defines a continuous map h : $\mathcal{W}\times(-\eta,\eta)^r\to(-\varepsilon',\varepsilon')^{d+r}$, such that $h(0)=0$, which is injective. By invariance of the domain, h is a homeomorphism onto some open neighborhood of 0, say $\mathcal{N}\subset(-\varepsilon',\varepsilon')^{d+r}$. For $(\widetilde{X},U)\in\mathcal{N}$, (F.27) can be inverted as

$$\overline{V}=\widetilde{\varphi}^3(\widetilde{X},U)-\widetilde{\varphi}^3(\widetilde{X},0),\tag{F.28}$$

and substituting (F.27) and (F.28) in (F.26) yields (F.8).

Finally we prove point 3, keeping in mind the previous definitions and properties of h, $\mathcal{W}$, η and $\mathcal{N}$. For $\widetilde{X}=(\widetilde{X}^1,\widetilde{X}^2)\in(-\varepsilon',\varepsilon')^d$, define $\overline{V}(\widetilde{X})\in I\!R^s\times\{0\}\subset I\!R^r$ by the formula :

$$J_r^s\overline{V}(\widetilde{X})=\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1,\widetilde{X}^2)\big(\widetilde{f}^2(0,0,0)-\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,0)\big). \tag{F.29}$$

Clearly $\overline{V}:(-\varepsilon',\varepsilon')^d\to I\!R^r$ is continuous and $\overline{V}(0)=0$, so there exists an open neighborhood $\mathcal{V}\subset\mathcal{W}$ of 0 in $I\!R^d$ such that $\overline{V}(\widetilde{X})\in(-\eta,\eta)^r$ as soon as $\widetilde{X}\in\mathcal{V}$; then, if we set $h(\widetilde{X},\overline{V}(\widetilde{X}))=(\widetilde{X},U(\widetilde{X}))\in\mathcal{N}$, it follows from (F.29), (F.28), and (F.8) that

$$\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,U(\widetilde{X}))=\widetilde{f}^2(0,0,0),\qquad \widetilde{X}\in\mathcal{V}. \tag{F.30}$$

We will show, using Proposition 3.9*, that the restriction of* $\widetilde{\varphi}_{\rm I}$ *to any relatively compact open subset* $\mathcal{X}$ *of* $\mathcal{V}$ *conjugates* (F.10) *and* (F.11) *over* $\mathcal{X}$*,* $\widetilde{\varphi}(\mathcal{X})$*, and this will achieve the proof.* To this effect, let $\mathcal{C}$ to be the collection of all piecewise affine maps $I\!R\to I\!R^s$ with constant slope $\widetilde{f}^2(0,0,0)$ (*cf* the discussion before Proposition 3.9) and note that, for any open set $\mathcal{O}\subset I\!R^s$ and any compact interval $J\subset I\!R$, the restriction of $\mathcal{C}$ to J contains, in its uniform closure, the set all piecewise continuous maps $J\to\mathcal{O}$. Now, consider a solution $\gamma\ :I\to\mathcal{V}$ of the control system :

$$\dot{\widetilde{X}}^1 \;=\; \widetilde{f}^1(\widetilde{X}^1,\Upsilon) \tag{F.31}$$

with state $\widetilde{X}^1$ and control Υ; hereafter, $\mathcal{V}_{\rm I}\subset I\!R^{d-s}$ and $\mathcal{V}_{\rm II}\subset I\!R^s$ will indicate the projections of $\mathcal{V}$ onto the first $d-s$ and the last s components respectively, and similarly for any other open set in $I\!R^d$. Assume that the control function $\gamma_{\rm II}\ :I\to\mathcal{V}_{\rm II}$ is the restriction to I of some member of $\mathcal{C}$. By definition, if a,b are the endpoints of I (that may belong to I or not), there are time instants $a=t_0<t_1<\cdots<t_N=b$, and vectors $\bar{\xi}_1,\ldots,\bar{\xi}_N\in I\!R^s$ such that, for $1\le j<N$, one has

$$t_{j-1}<t<t_j \;\;\Rightarrow\;\; \gamma_{\rm II}(t) \;=\; \bar{\xi}_j+t\widetilde{f}^2(0,0,0), \tag{F.32}$$

while at the points t_j themselves $\gamma_{\rm II}$ is either right or left continuous when $1<j<N$. *We claim* that $\widetilde{\varphi}_{\rm I}(\gamma(t))$ is a solution that remains in $\widetilde{\varphi}_{\rm I}(\mathcal{V})$ of the control system :

$$\dot{Z}^1 \;=\; g^1(Z^1,\Gamma) \tag{F.33}$$

with state Z^1 and control Γ. In fact, since $\gamma_{\rm I}$ is continuous by definition of a solution, so is $\widetilde{\varphi}^1(\gamma_{\rm I})$ and therefore, as $\widetilde{\varphi}_{\rm I}(\gamma(t))$ lies in $\widetilde{\varphi}_{\rm I}(\mathcal{V})$ for all $t\in I$ by construction, it is enough to check that

$$\widetilde{\varphi}^1(\gamma_{\rm I}(T_2))-\widetilde{\varphi}^1(\gamma_{\rm I}(T_1))=\int_{T_1}^{T_2}g^1\big(\widetilde{\varphi}^1(\gamma_{\rm I}(t)),\widetilde{\varphi}^2(\gamma_{\rm I}(t),\gamma_{\rm II}(t))\big)\,\mathrm{d}t \tag{F.34}$$

whenever $t_{j-1}<T_1<T_2<t_j$ for some $j>1$. However, the restriction of $\gamma(t)$ to (t_{j-1},t_j) is a solution that remains in $\mathcal{V}$ of the differential equation :

$$\begin{aligned}\dot{\gamma}_{\rm I} &= \widetilde{f}^1(\gamma_{\rm I},\gamma_{\rm II})\\ \dot{\gamma}_{\rm II} &= \widetilde{f}^2(0,0,0),\end{aligned}$$

hence $(\gamma(t), U(\gamma(t))$ is, by (F.30), a solution of (F.6) that remains in $\mathcal{N}$, and therefore (F.34) follows from the triangular structure (F.7) of $\widetilde{\varphi}$ and the fact that it conjugates system (F.6) to system (F.5). *This proves the claim.*

In the other direction, we observe since it is included in $\mathcal{W}$ that $\mathcal{V}$ has compact closure in $(-\varepsilon', \varepsilon')^d$, and therefore that $\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})$ in turn has compact closure in $\widetilde{\varphi}_{\mathrm{I}}\big((-\varepsilon', \varepsilon')^d\big)$. Pick $\eta' > 0$ such that $\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V}) \times (-\eta', \eta')^r \subset \widetilde{\varphi}((-\varepsilon', \varepsilon')^{d+r})$, and let $\mathcal{C}'$ denote the collection of all piecewise smooth maps $\mathbb{R} \to \mathbb{R}^s$ whose derivative is strictly bounded by η' component-wise. The restriction of $\mathcal{C}'$ to any compact real interval J is uniformly dense in the set all piecewise continuous maps $J \to \mathcal{O}$, for any open set $\mathcal{O} \subset \mathbb{R}^s$. Clearly, any solution $\gamma' : I \to \widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})$ of system (F.33), whose control function $\gamma'_{\mathrm{II}} : I \to \big(\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})\big)_{\mathrm{II}}$ is the restriction to I of some member of $\mathcal{C}'$, satisfies the differential equation

$$\begin{array}{rcl} \dot{\gamma}'_{\mathrm{I}} & = & g^1(\gamma'_{\mathrm{I}}, \gamma'_{\mathrm{II}}) \\ \dot{\gamma}'_{\mathrm{II}} & = & J^s_r\,(d\gamma'_{\mathrm{II}}/dt\,,\,0) \end{array}$$

on every interval where it is smooth. By the very definition of η' and $\mathcal{C}'$, it follows that $\big(\gamma'(t), (d\gamma'_{\mathrm{II}}(t)/dt, 0)\big)$ is, on such intervals, a solution to (F.5) that remains in $\widetilde{\varphi}((-\varepsilon', \varepsilon'))^{d+r}$ and, since $\widetilde{\psi}$ conjugates system (F.5) to system (F.6), we argue as before to the effect that $\widetilde{\psi}_{\mathrm{I}}(\gamma')$ is a solution to system (F.31) that remains in $\mathcal{V}$. Appealing to Proposition 3.9, we conclude that $\widetilde{\varphi}_{\mathrm{I}}$ conjugates system (F.31) to system (F.33) on relatively compact open subsets of $\mathcal{V}$, as desired. $\square$

Acknowledgements

The authors wish to thank Prof. I. Kupka, from Université Pierre et Marie Curie (Paris VI) for many enriching discussions and for pointing out reference [6] to them. They also want to acknowledge discussions with Prof. F. Colonius from Universität Augsburg and Prof. W. Kliemann from Iowa State University on their work on flows for control systems (see section 7.3) and on Grobman-Hartman theorems in this setting. These contacts were initiated within the European Community TMR Program "Nonlinear Control Network", # ERB FMRXCT-970137. Finally, thanks are also due to Prof. C.T.C. Wall from the University of Liverpool for his comments on the open question of section 5.2.

References

[1] V. I. Arnold, *Equations Différentielles Ordinaires*, MIR, 3 ed., 1974.

[2] P. Brunovský, *A classification of linear controllable systems*, Kybernetika, 6 (1970), pp. 176–188.

[3] S. Celikovský, *Topological equivalence and topological linearization of controlled dynamical systems*, Kybernetika, 31 (1995), pp. 141–150.

[4] E. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, 1955.

[5] F. Colonius and W. Kliemann, *Some aspects of control systems as dynamical systems*, J. Dyn. Differ. Equations, 5 (1993), pp. 469–494.

[6] N. D. Cong, *Topological dynamics of random dynamical systems*, Clarendon Press, Oxford Science Publ., 1997.

[7] J. Dieudonné, *Eléments d'Analyse*, volume 1, Gauthier-Villars, 1969.

[8] P. Hartman, *Ordinary Differential Equations*, Birkhäuser, 2 ed., 1982.

[9] L. R. Hunt, R. Su, and G. Meyer, *Design for multi-input nonlinear systems*, in Differential Geometric Control Theory, R. Brockett, ed., Birkhäuser, 1983, pp. 258–298.

[10] A. Isidori, *Nonlinear Control Systems*, Comm. in Control Engineering, Springer-Verlag, 3rd ed., 1995.

[11] B. Jakubczyk, *Equivalence and invariants of nonlinear control systems*, in Nonlinear controllability and optimal control, Dekker, New York, 1990, pp. 177–218.

[12] B. Jakubczyk and W. Respondek, *On linearization of control systems*, Bull. Acad. Polon. Sci. Sér. Sci. Math., 28 (1980), pp. 517–522.

[13] V. JURDJEVIC, *Geometric Control Theory*, vol. 51 of Cambridge Studies in Advanced Mathematics, Cambridge Univ. Press, 1997.

[14] V. JURDJEVIC AND H. J. SUSSMANN, *Controllability of nonlinear systems*, J. of Diff. Equations, 12 (1972), pp. 95–116.

[15] C. LOBRY, *Contrôlabilité des systèmes non linéaires*, SIAM J. on Control, 8 (1970), pp. 573–605.

[16] J. R. MUNKRES, *Elements of Algebraic Topology*, Addison-Wesley, 1984.

[17] T. NAGANO, *Linear differential systems with singularities and an application to transitive lie algebras*, J. Math. Soc. Jap., 18 (1966), pp. 398–404.

[18] H. NIJMEIJER AND A. J. VAN DER SCHAFT, *Nonlinear dynamical control systems.*, Springer-Verlag, New York, 1990.

[19] R. ROUSSARIE, *Modèles locaux de champs et de formes*, vol. 30 of Astérisque, Soc. Math. de France, 1975.

[20] W. RUDIN, *Functional analysis*, Mc Graw Hill, 1973.

[21] W. RUDIN, *Analyse réelle et complexe*, Masson, Paris, 1975.

[22] E. D. SONTAG, *Mathematical Control Theory*, Springer Texts in Applied Mathematics, 6, 1998.

[23] M. SPIVAK, *A Comprehensive Introduction to Differential Geometry*, vol. 1, Publish or Perish, Houston (U.S.A.), second ed., 1979.

[24] H. J. SUSSMANN, *Orbits of families of vector fields and integrability of distributions*, Transactions of the AMS, 317 (1973), pp. 171–188.

[25] K. TCHOŃ, *The only stable normal forms of affine systems under feedback are linear*, Syst. & Control Lett., 8 (1987), pp. 359–365.

[26] A. J. VAN DER SCHAFT, *Linearization and input-output decoupling for general nonlinear systems.*, Syst. & Control Lett., 5 (1984), pp. 27–33.

[27] J. C. WILLEMS, *Topological classification and structural stability of linear systems*, J. of Diff. Equations, 35 (1980), pp. 306–318.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399